

John Rawls, recientemente fallecido en Estados Unidos, es considerado el filósofo político más importante del siglo XX. Su **Teoría de la justicia** ha vendido centenares de miles de ejemplares en todo el mundo. La pregunta central que orientó el trabajo de Rawls es tan vieja como la filosofía misma: **¿qué es una sociedad justa?** O, de manera más precisa: ¿cuáles son los principios fundamentales sobre los que se deberían edificar sus instituciones, y en función de los cuáles deberíamos evaluar las políticas desarrolladas por los gobiernos?

Liberal y al mismo tiempo igualitarista, Rawls discutió a derecha y a izquierda, con utilitaristas, libertaristas, igualitaristas radicales y comunitaristas, y terminó por convertirse en una referencia ineludible para el pensamiento político contemporáneo. Este libro repasa su trayectoria y explica el significado de sus diferentes obras.

campo de ideas


ISBN 84-96089-07-X
9 788496 089075

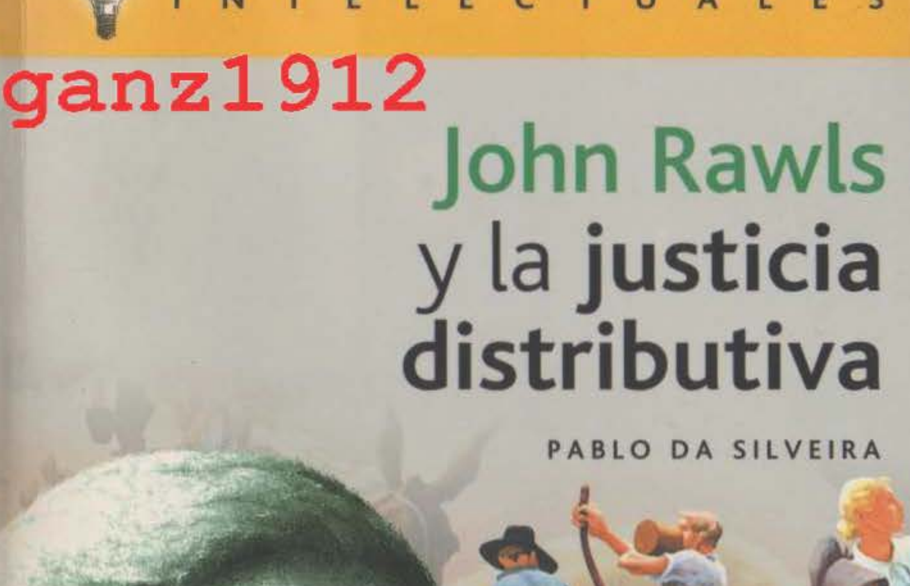
JOHN RAWLS Y LA JUSTICIA DISTRIBUTIVA

i

INTELECTUALES

John Rawls y la justicia distributiva

PABLO DA SILVEIRA



I N T E L E C T U A L E S

John Rawls y la justicia distributiva

PABLO DA SILVEIRA

campo de ideas

JOHN RAWLS Y LA JUSTICIA DISTRIBUTIVA

Directores de la serie Intelectuales:
J.C. Kreimer, M.J. Stuart
intelectuales@ciudad.com.ar

Investigación y texto: Pablo da Silveira
Corrección: Taller de Publicaciones SLL, Madrid.
Diseño gráfico: txt ediciones

Campo de Ideas SL,
Calle Alquimia, 6
28933 Móstoles (Madrid) España
Tel. 91 617 08 67 - Fax: 91 617 97 14
E-mail: editorial@alfaomega.es

Primera edición: abril 2003
Depósito Legal: M. 17.062-2003
I.S.B.N.: 84-96089-07-x
Impreso en España por Artes Gráficas Cofás, S.A.

Índice

INTRODUCCIÓN

Formas de desigualdad

¿Qué es más importante? ¿Que todos tengamos lo mismo o que todos vivamos mejor? A primera vista puede parecer que se trata de la misma cosa, pero basta con pensar un poco para darse cuenta de que no. Hagamos un experimento mental. Imaginemos una sociedad compuesta por sólo dos personas a las que vamos a llamar el individuo A y el individuo B (es obvio que no existen sociedades tan pequeñas, pero esta clase de simplificación ayuda a ver más claro). Ahora imaginemos dos situaciones posibles. En la primera, tanto A como B tienen 2 unidades de algo que se pueda repartir fácilmente. No importa demasiado de qué se trate. Pueden ser dólares, chocolates o cebollas. Lo verdaderamente importante es que es mejor tener una mayor cantidad de esas unidades que tener menos, porque eso nos pone en mejores condiciones de satisfacer nuestras preferencias. Dicho de un modo ligeramente más técnico: estas unidades pueden ser consideradas un indicador razonablemente confiable de nuestro nivel de bienestar.

En la primera situación, entonces, tanto el individuo A como el individuo B tienen 2 unidades. La cantidad total (lo que los economistas llamarían el producto) es 4. Podemos representar este estado de cosas de la siguiente manera:

A	2
B	2
	4

Esta es una sociedad perfectamente igualitaria. Todos sus miembros tienen lo mismo. Pero si se diera el caso de que esas dos unidades fueran muy poco (por ejemplo, podría tratarse de 2.000 dólares para vivir durante todo un año) sería también una sociedad pobre. Los individuos A y B tendrían la misma cantidad de unidades pero sólo podrían satisfacer una pequeña parte de sus preferencias.

Imaginemos ahora una segunda situación. En este nuevo estado de cosas, el individuo A tiene 6 unidades y el individuo B tiene 10, lo que arroja un total de 16. La situación puede representarse así:

A	6
B	10
	16

Si pudiéramos pasar a esta segunda situación, tendríamos una sociedad bastante más rica que en el caso anterior.

El producto habría pasado de 4 a 16, lo que quiere decir que se habría cuadriplicado. Tendríamos también una situación menos igualitaria, ya que en este caso el individuo B tendría bastantes más unidades que el individuo A. Sin embargo, también el individuo A se habría beneficiado. Sin duda sería ahora menos rico que B, pero sería tres veces más rico de lo que él mismo era en la primera situación (para seguir con el ejemplo: ahora tendría 6.000 dólares para vivir todo un año, en lugar de 2.000).

Gesto de valentía. El primero de John Rawls consiste en aceptar que algo así puede suceder: **cierto grado de desigualdad en la distribución de la riqueza puede ser necesario para permitir que todos vivamos mejor.** Por ejemplo, podría ocurrir que, gracias a que el individuo B tiene ahora 10 unidades, esté en condiciones de invertir y darle empleo al individuo A. Si así fueran las cosas, la desaparición de la igualdad estricta y la mejora de la calidad de vida de todos los miembros de la sociedad estarían estrechamente relacionadas.

Es probable que esta conclusión sea difícil de aceptar para quienes tienen una sensibilidad igualitarista y que al mismo tiempo deje satisfechos a quienes piensan que la igualdad es poco importante. Pero estos estados de ánimo no durarían mucho tiempo, porque Rawls agrega enseguida que este razonamiento no debe convertirse en una excusa para aceptar cualquier forma de desigualdad. Si bien algunas formas de distribución desigual pueden ser

necesarias para permitir que todos vivamos mejor, de aquí no se sigue que toda forma de desigualdad deba ser aceptada. Puede que haya formas de desigualdad inevitables pero hay muchas otras que sencillamente debemos rechazar. Esta afirmación rotunda es el segundo gesto de valentía de Rawls. Y si el primero había hecho sonreír a quienes piensan que la igualdad no es importante, el segundo hará sonreír a los igualitaristas. Porque Rawls es de los que piensa que el Estado debe distribuir.

La pregunta que debemos hacernos no es, entonces, si debemos aceptar o rechazar la igualdad estricta, sino otras más sofisticadas: ¿cuáles son las formas de desigualdad que podemos considerar aceptables? ¿Y cuáles son las formas de desigualdad que en ningún caso deberíamos admitir? John Rawls es un filósofo contemporáneo que dedicó su vida a intentar resolver estos interrogantes.

¿Quién es John Rawls?

John Rawls nació en Baltimore el 21 de febrero de 1921. Hizo sus estudios secundarios en Connecticut y se graduó en la Universidad de Princeton en 1943. La Segunda Guerra Mundial estaba en su apogeo en ese entonces, por lo que fue movilizado de inmediato y participó hasta 1946 en la campaña del Pacífico. Estuvo en Filipinas, Nueva Guinea y Japón. Más tarde dirá que esa experiencia le convenció para siempre de «la enorme importancia de que los militares estén subordinados al gobierno civil».

En 1949 se casó con Margaret Fox y en 1950 obtuvo su doctorado. A partir de allí inicia lo que prometía ser la tranquila vida de un profesor universitario en un país rico: primero trabaja como instructor en Princeton, unos años más tarde se traslada a Cornell, donde obtendrá un puesto de profesor titular en 1962, luego pasará por el Instituto Tecnológico de Massachusetts (MIT) y finalmente se instalará en Harvard, donde impartirá la docencia hasta el momento de retirarse.

Desde mediados de los años 50, Rawls ganó cierto prestigio como estudioso de un conjunto de problemas que se sitúan en la frontera entre la ética y la economía. En esa época publicó algunos artículos que fueron bien recibidos en los círculos de especialistas. Pero, si bien esos textos resultan significativos a la luz de lo que pasó después, para el gran público seguía siendo un desconocido. Cuando estaba a punto de cumplir los cincuenta, sólo sus alumnos de Harvard y un puñado de académicos profesionales tenían noticia de su existencia.

Obras. Así estaban las cosas cuando, en 1971, Rawls publicó un libro que había estado preparando durante años. Su título era *A Theory of Justice,* es decir, "Una teoría de la justicia" (aunque la traducción al castellano suena más contundente al eliminar la primera palabra). Esa obra relativamente tardía tuvo un éxito inmediato. Primero en Estados Unidos, luego en Inglaterra y en otros países anglófonos, poco más tarde en el resto de Europa y finalmente en todo

el mundo, una enorme cantidad de filósofos profesionales, economistas y lectores interesados en la justicia social la colocaron en el centro de sus discusiones. Hasta tal punto que, en el año 1974, el filósofo estadounidense Robert Nozick (que discrepaba con Rawls en todo lo que dos filósofos pueden discrepar) escribió la siguiente frase: «De ahora en adelante los filósofos políticos deben trabajar dentro de la teoría de Rawls, o bien deben explicar por qué no lo hacen.» (Nozick, 1974, pág. 183). Los editores del libro no podían salir de su asombro: un texto técnico y bastante árido, sin ningún encanto literario ni referencias a cuestiones de actualidad, acababa de convertirse en un *best-seller*.

Durante los años setenta y ochenta, Rawls se vio involucrado en una gran cantidad de debates con autores que criticaban o intentaban modificar su teoría. Pero durante todos esos años no publicó un solo libro. Las respuestas que daba a sus críticos se acumularon bajo la forma de artículos editados en revistas académicas. Esos artículos más o menos dispersos atrajeron la atención de muchos lectores y dieron lugar a una inmensa bibliografía secundaria.

1993. Así fueron las cosas hasta que en ese año Rawls publicó el segundo de sus libros: *El liberalismo político* (que fue rápidamente traducido al castellano, esta vez sin cambios en el título). La nueva obra demostraba que su autor había sido muy sensible a algunas de las críticas que había recibido, hasta el punto de haber introducido modificaciones importantes en sus ideas y en su estilo de trabajo. Tal co-

mo había ocurrido con Wittgenstein, los comentaristas empezaron a hablar de un «primer» y de un «segundo Rawls».

Teoría de la justicia había llamado la atención por su sólida arquitectura conceptual, por su argumentación despojada y elegante, y por el carácter abstracto y universalista de sus conclusiones. *El liberalismo político* es muy distinto en todos estos aspectos. En primer lugar, es un libro que carece de una arquitectura clara, hasta el punto de que Rawls lo presenta como una serie de conferencias. En segundo lugar, la argumentación es menos despojada (muchos dicen que menos elegante) y más abierta a la consideración de problemas particulares. Por último, la intención inicial de pensar principios de justicia que pudieran ser aceptados por todos los individuos racionales parece haber quedado a un lado.

El segundo libro de Rawls abrió una nueva ronda de discusiones, pero eso no le impidió seguir adelante con nuevos proyectos. En 1993 (el mismo año que publicó *El liberalismo político)* fue invitado por Amnistía Internacional a dictar una conferencia sobre derechos humanos. En esa exposición, Rawls se propuso hacer lo mismo que Kant había hecho dos siglos atrás: sacar su pensamiento político de los límites de una sociedad específica y explorar sus posibles consecuencias a escala mundial.

El texto de esa conferencia fue publicado ese mismo año y tuvo el efecto de involucrar en el debate a los especialistas en relaciones internacionales.

Dos años más tarde, Rawls se embarcó en una discusión por escrito con Jürgen Habermas, que fue inicialmente publicada en una prestigiosa revista académica *(The Journal of Philosophy)* y luego editada como libro en una gran cantidad de idiomas. La razón es fácil de entender: los dos filósofos políticos más influyentes de la segunda mitad del siglo XX finalmente habían decidido confrontar sus puntos de vista.

En 1995 Rawls sufrió el primero de una larga serie de accidentes circulatorios que lo obligaron a abandonar progresivamente la actividad. Pero, curiosamente, después de esa fecha se han publicado más páginas con su firma que en las décadas anteriores. En 1996 apareció la edición en rústica de *El liberalismo político,* para la que Rawls escribió una importante introducción. En 1999 editó *El derecho de gentes,* que incluye entre otros materiales una versión corregida y ampliada de la conferencia de Amnistía. Ese mismo año Harvard University Press publicó un grueso volumen titulado *Collected Papers,* que reúne casi todos los artículos publicados por Rawls a lo largo de su vida. La obra fue preparada por Samuel Freeman, un rawlsiano convencido que trabajó durante mucho tiempo en el proyecto bajo la supervisión del propio Rawls.

Dos nuevas obras. Se sumaron recientemente a su obra que reúnen materiales preparados por otras personas. En el año 2000 apareció un libro editado por Barbara Herman que recoge algunos de los textos que produjo Rawls (o que

reconstruyeron sus alumnos) durante sus años de profesor en Harvard. La obra ayuda a entender la génesis de su pensamiento y permite conocer la profundidad de su erudición filosófica, normalmente oculta en sus textos sistemáticos. Al año siguiente, Erin Kelly editó en forma de libro un texto legendario que durante años había circulado por las universidades del mundo en versiones fotocopiadas. Se trata de un intento de reformulación de la teoría de 1971, que a juicio del propio Rawls nunca alcanzó el grado de madurez suficiente como para merecer convertirse en libro, pero que, en opinión de muchos, contiene algunas de las mejores páginas escritas por su autor.

Rawls pasó los últimos años de su vida alejado de los foros filosóficos. Su obra seguía ganando influencia con el paso del tiempo, pero su imagen personal empezaba a volverse un recuerdo. Alto, delgado, sin grandes dotes de orador pero agradable y reposado en el diálogo, siempre dio la impresión de ser al mismo tiempo una persona extremadamente inteligente y un hombre bueno (dos cosas que no siempre van juntas, por cierto). Murió a los ochenta y dos años, el 25 de noviembre de 2002.

CAPÍTULO 1

Un estallido en el mundo anglosajón

En general todos pensamos que la justicia es importante. Todos preferimos vivir en una sociedad justa antes que vivir en una sociedad donde reine la injusticia. Pero ¿cómo saber cuando estamos ante una sociedad justa? O dados dos estados posibles en los que puede encontrarse una sociedad, ¿cómo decidir cuál de ellos es el más justo?

Estas preguntas son tan viejas como la filosofía misma. Hace unos dos mil quinientos años, Platón escribió una larga obra *(La República)* en la que intentaba darles respuesta. Pero los siglos se han sucedido sin que hayamos conseguido llegar a conclusiones definitivas, lo que llevó a mucha gente a pensar que simplemente no estamos equipados para resolver esta clase de problemas. Podemos calcular con precisión la velocidad a la que viaja la luz, pero no podemos construir argumentos sobre la justicia (ni sobre lo bueno, ni sobre lo moralmente correcto) que puedan aspirar a ser aceptados por un auditorio más o menos universal.

Analíticos. En la época en la que Rawls se formó como filósofo, este escepticismo era la ortodoxia en el mundo anglosajón. El método de trabajo predominante entre los filósofos profesionales era el llamado «método analítico», que había renunciado a dar respuestas sustantivas a esta clase de preguntas. Los filósofos analíticos pensaban que la mayor parte de los problemas de los que tradicionalmente se ha ocupado la filosofía son en realidad confusiones lingüísticas. No se trata de problemas genuinos, sino de simples sinsentidos producidos por un uso impreciso del lenguaje. Por tanto, la única tarea que todavía pueden cumplir los filósofos consiste en aclarar los términos utilizados en las discusiones, así como en criticar aquellos usos que sean inconsistentes o imprecisos[1].

Los filósofos analíticos. Por cierto, no estaban de acuerdo a la hora de precisar el alcance de este programa. Algunos pensaban que *todo* el trabajo debe reducirse al análisis lingüístico, ya que la discusión sustantiva es o bien innecesaria (porque los problemas desaparecen una vez que se aclaran los términos), o bien inconducente (porque no hay manera de zanjarla mediante procedimientos racionales). Otros veían al esfuerzo analítico como una etapa preparatoria de la discusión sustantiva. Pero, fuera cual fuera la opinión de cada uno, la gran mayoría dedicaba mucho más

1 Sigo la presentación realizada en Da Silveira & Norman, 1995.

tiempo y energía a las clarificaciones lingüísticas que a la discusión sustantiva.

Esta era la ortodoxia en los tiempos en los que se formó Rawls, pero no es este el punto de vista que encontramos en *Teoría de la justicia*. En ese libro Rawls muestra mucho respeto por los filósofos analíticos, en el sentido de que se toma muy en serio sus exigencias de precisión lingüística y claridad conceptual. Pero al mismo tiempo entra en conflicto con ellos al negar que la reflexión filosófica deba limitarse a realizar clarificaciones lingüísticas. No todos los problemas morales ni todos los problemas políticos son el resultado de un uso impreciso del lenguaje. Existen situaciones reales a las que hay que dar respuestas reales, y esas respuestas pueden ser mejores o peores desde el punto de vista normativo (por ejemplo, pueden ser más o menos justas). El desafío consiste en desarrollar argumentos que nos permitan optar racionalmente entre esas soluciones.

Un filósofo posanalítico. Al asumir esta postura, Rawls estaba dando un paso decisivo en el camino que conduce desde la filosofía analítica a lo que hoy suele llamarse la filosofía posanalítica. El prefijo «pos» alude a algo más que una mera sucesión temporal. Para los filósofos posanalíticos, las exigencias de precisión y de claridad siguen siendo importantes, pero lo que se niega es que la tarea del filósofo deba reducirse a un mero «pulir las herramientas». El mundo de la ética y de la política nos exigen tomar decisiones, y para fundamentarlas necesitamos argumentos sus-

tantivos. Lo que se espera del filósofo es que haga esfuerzos para encontrarlos.

En resumen, Rawls había decidido volver a ocuparse de algunos de los problemas de fondo de los que antes se había ocupado la gran tradición filosófica. Pero con esto no hacía más que meterse en problemas, porque simultáneamente tenía claro que no podía servirse de los mismos argumentos que habían usado los filósofos de los siglos anteriores. Este es un aspecto esencial del desafío filosófico al que se enfrentó, de modo que conviene prestarle atención.

Antifundacionalismo. Cuando John Locke tuvo que explicar por qué los seres humanos tenemos derecho a la vida y a la integridad física, apeló como fundamento a nuestra condición de hijos de Dios. Cuando los filósofos de la Ilustración francesa afirmaron que los seres humanos tenemos derechos inalienables, argumentaron que esos derechos forman parte de nuestra común naturaleza humana y que esa naturaleza puede ser conocida mediante un uso adecuado de la razón. Cuando Hegel desarrolló su teoría del Estado moderno, se apoyó en la idea de que la historia avanza en un sentido determinado y que el Estado moderno debe ser comprendido como un estadio fundamental de ese avance.

El problema es que todas estas justificaciones se han vuelto problemáticas en las sociedades democráticas contemporáneas. Si la única razón para respetar la vida y la in-

tegridad física de los demás es aceptar nuestra condición de hijos de Dios, muchos de nuestros conciudadanos no tendrán motivos para respetar la vida ni la integridad física de nadie, por la sencilla razón de que no creen en Dios. Si la condición para aceptar la idea de derechos inalienables es admitir que existe una naturaleza humana y que nuestra razón está en condiciones de conocerla, no serán muchos los que tengan motivos para respetar esos derechos. Las teorías de la naturaleza humana se han desacreditado a los ojos de mucha gente porque cada vez que un teórico de esa corriente ha querido ofrecernos un retrato de nuestra naturaleza tal como la razón puede conocerla, nos ha ofrecido un retrato diferente del que ofrecían otros teóricos igualmente confiados en nuestras capacidades racionales. Y lo mismo ocurre con las ideas hegelianas sobre el sentido de la historia. Para muchos de nuestros contemporáneos, Hegel carece de argumentos sólidos para sostener que la historia tenga un sentido. Tal vez Shakespeare tenía razón cuando escribía en *Macbeth* que «la vida es un cuento narrado por un idiota, lleno de sonido y de furia. No significa nada».

El pluralismo. El dato fundamental a partir del que tenemos que intentar filosofar, dice Rawls, es lo que él mismo llamará más tarde «el *hecho del pluralismo*». Este hecho consiste en la radical diversidad de convicciones morales, metafísicas y religiosas con las que se identifican los integrantes de las sociedades democráticas contemporáneas. Según Rawls, tene-

mos que tener dos cosas muy claras respecto de esta situación. La primera es que estamos ante un hecho y no ante una opinión. Nos guste o no nos guste, las cosas son así. La segunda es que este hecho no tiene nada de pasajero. La diversidad que caracteriza a nuestras sociedades no es el tipo de desacuerdo que pueda superarse mediante la argumentación, sino el resultado del libre intercambio de ideas y del respeto hacia múltiples formas de vida. Esto tiene una consecuencia enormemente importante, y es que no podemos esperar a superar la diversidad de convicciones para embarcarnos en la tarea de construir instituciones comunes. El desafío consiste en diseñar instituciones políticas que puedan ser reconocidas como legítimas en un contexto de diversidad.

Este punto de vista, que se conoce con el nombre de «antifundacionismo», no nos obliga a dejar de lado nada en lo que creamos. No se trata de renunciar a nuestras convicciones profundas ni de renunciar a justificarlas con razones. A lo que hay que renunciar es a edificar directamente sobre ellas la justificación de las instituciones comunes.

El antifundacionismo supone reconocer que, aun cuando tengamos nuestras propias convicciones y aun cuando podamos invocar razones en su favor, esto no será suficiente para lograr la unanimidad en torno a ellas. Los otros pueden tener sus propias razones para defender convicciones más o menos incompatibles con las nuestras.

Rawls quedaba así enfrentado a un doble desafío. Quería ir más allá del límite que se habían autoimpuesto los

filósofos analíticos, porque consideraba esencial que pudiéramos dar respuestas sustantivas a nuestras preguntas sobre la justicia. Pero al mismo tiempo renuncia a los argumentos que habían sido tradicionalmente utilizados para resolver estas cuestiones, es decir, aquellos fundados en convicciones controvertidas acerca de la voluntad de Dios, nuestra propia naturaleza o el sentido de la historia. ¿Es posible encontrar una solución a un problema planteado en estos términos?

Una metáfora famosa

La solución que propuso Rawls es en cierto sentido sorprendente, porque suponía retomar contacto con una tradición intelectual que parecía haber caído en desuso. Se trata de la vieja tradición del contrato social.

El contractualismo había sido impulsado en los siglos XVII y XVIII por autores tan importantes como Thomas Hobbes, John Locke, Jean-Jacques Rousseau e Immanuel Kant. Cada uno de ellos lo había adaptado a sus intereses filosóficos y cada uno había llegado a conclusiones diferentes a las defendidas por los demás. Todos compartían, sin embargo, una idea esencial: la mejor manera de pensar las instituciones que una sociedad debe darse consiste en imaginar que esa sociedad es el fruto de un contrato voluntario entre sus miembros, y luego preguntarse si una institución específica podría o no formar parte de ese contrato. Si la respuesta es afirmativa, podemos considerar que esa

institución es legítima. Si la respuesta es negativa, debe ser rechazada o profundamente reformada.

Los contractualistas importantes creen que la idea del contrato es un artificio que ayuda a pensar. Ninguno de ellos pensó que un acuerdo de ese tipo hubiera tenido lugar. En el fondo, se trataba de algo similar a los modelos que hoy utilizan los economistas. Y, como ocurre con los modelos económicos, los resultados a los que se llegaba dependían en gran medida del modo en que los individuos eran descritos en el momento inicial. Si concebimos a los individuos como puramente movidos por intereses egoístas (como hacía Hobbes) vamos a llegar a cierto tipo de conclusiones. Si los vemos como capaces de benevolencia y de cooperación (como suponía Rousseau) vamos a llegar a resultados diferentes. Dentro de la tradición contractualista, el modo en que se procesó la discusión sobre los supuestos del modelo consistió en definir las características de lo que se llamaba «el estado de naturaleza».

Esta clase de discusión apasionó a los filósofos políticos durante mucho tiempo, pero en el momento en que apareció *Teoría de la justicia* era considerada cosa del pasado. El contractualismo había alcanzado su culminación en el siglo XVIII y luego había pasado a integrar la galería de las venerables inutilidades. Ciertamente se seguía enseñando en las facultades de filosofía, como se enseñan muchas otras ideas que tuvieron cierta importancia en algún momento de la historia. También era verdad que los principales con-

tractualistas habían tenido una fuerte influencia sobre el pensamiento constitucional posterior. Pero a lo largo de todo el siglo XIX y durante los tres primeros cuartos del siglo XX nadie esperaba que se produjeran nuevos desarrollos filosóficos a partir de la vieja lógica contractual. Hasta que llegó John Rawls.

Por cierto, lo que Rawls propuso no fue un simple retorno nostálgico a los tiempos de Rousseau. Su propuesta tenía toda la sofisticación y todo el despojamiento esperables en un filósofo que había estado expuesto a la tradición analítica. Pero la metáfora de la que partía era exactamente la misma que habían utilizado los contractualistas varios siglos antes.

Posición Original. Imaginemos una asamblea en la que todos los miembros de una sociedad van a elegir sus instituciones de base. Por «instituciones de base» se entiende el conjunto de instituciones fundamentales que determinarán el modo en que se organizará la vida económica y política de esa sociedad. Rawls llama a esta asamblea imaginaria la *posición original*. Los individuos que participan en ella saben que las decisiones que tomen van a condicionar toda su vida, así como la vida de sus hijos y de sus nietos. No se trata de tomar medidas coyunturales, sino de fijar reglas de juego muy estables. Todos aceptan que esas reglas serán fijadas mediante un acuerdo que contemple los intereses y preocupaciones de cada uno.

Agreguemos dos nuevos datos. El primero es que, por más diferentes que sean entre sí, todos los miembros de la sociedad comparten dos *capacidades morales fundamentales*: la capacidad de desarrollar y de intentar poner en práctica una concepción del bien (esto es, un conjunto de ideas acerca de cómo vale la pena vivir que orientará sus elecciones particulares) y la capacidad de desarrollar un sentido de la justicia (esto es, la capacidad de entender los principios normativos que se haya decidido respetar y de actuar en consonancia con ellos). Rawls resumirá estas ideas diciendo que los individuos son al mismo tiempo *racionales* y *razonables*.

El segundo dato es que, en el momento de ingresar en esta asamblea imaginaria, los individuos caen bajo un *velo de ignorancia*. Esto significa que los individuos «olvidan» cuáles son sus características particulares: no saben si son ricos o pobres, hombres o mujeres, blancos o negros, creyentes o no creyentes. Tampoco conocen cuál es su propia concepción del bien, es decir, el «sistema de fines» que va a orientar sus elecciones.

Este *velo de ignorancia* no implica que los individuos lo olviden todo. De hecho, los participantes en la *posición original* siguen disponiendo de abundante información. Saben que hay negros y blancos, mujeres y hombres, creyentes y no creyentes. Saben que tienen discrepancias a propósito de cómo vale la pena vivir y saben que son capaces de ajustar su comportamiento a exigencias normativas. También sa-

ben que en la sociedad en la que viven hay escasez de recursos, conocen las leyes fundamentales de la economía, y las condiciones culturales y sociales que deben darse para asegurar el buen funcionamiento de las instituciones. Lo único que no conocen es *su propio lugar* en la sociedad, ni cuáles son sus ideas acerca de cómo vale la pena vivir. Se cuenta con conocimientos generales pero no se tiene ninguna información particular.

El velo de ignorancia es la gran diferencia entre el contractualismo de Rawls y el de Thomas Hobbes. El contractualismo hobbesiano está basado en la idea de mutuo beneficio. Los individuos se ponen de acuerdo en una serie de normas porque esto favorece sus intereses particulares. En el caso de Rawls, los intereses particulares son justamente lo que se ignora: los participantes en la *posición original* deben llegar a acuerdos sin conocer lo que les beneficia ni lo que les perjudica.

La pregunta que surge de inmediato es por qué los individuos deberían aceptar semejante restricción. Y Rawls responde diciendo que hacerlo es una decisión racional. Recordemos que las instituciones de base que sean pactadas en la *posición original* van a determinar las condiciones en las que se va a desarrollar nuestra vida, la de nuestros hijos y la de nuestros nietos. Cuando uno toma decisiones de tan largo alcance debe tener presente que las circunstancias pueden cambiar enormemente a lo largo del tiempo. Puede que hoy seamos ricos, pero nada asegura que dentro de

unos años no seamos pobres. Puede que ahora estemos en mayoría, pero nada impide que mañana quedemos en minoría. Por tanto, los términos del contrato no deben ser evaluados a la luz de nuestros intereses actuales, sino de los intereses más permanentes que podamos tener. La vida social está demasiado cargada de incertidumbre como para que sea racional tomar decisiones en función de las ventajas de las que disfrutamos en el momento presente.

Aceptar el *velo de ignorancia* va entonces en favor de nuestros propios intereses a largo plazo. Pero eso no quita que, una vez que lo hemos adoptado, la lógica del propio interés queda neutralizada. En definitiva, lo que hace el *velo de ignorancia* es obligarnos a negociar bajo la perspectiva del universalismo moral. Detrás de toda la argumentación de Rawls está la expectativa optimista de una confluencia final entre racionalidad (en el sentido de autointerés) y moralidad.

Decidiendo en la *posición original*

Al igual que todos los grandes contractualistas, Rawls niega que una asamblea de este tipo haya tenido lugar[2]. También afirma que el retrato simplificado de nosotros mismos no es una generalización empírica, sino que los individuos están «definidos teóricamente». Pero eso no le impide creer que la *posición original* resume de manera adecua-

2 *Véase*, por ejemplo, el párrafo 20 de *Theory*.

da los datos esenciales en los que se desarrolla la vida social (la escasez de recursos, la multiplicidad de concepciones del bien), así como las características esenciales de cada uno de nosotros en tanto que participantes de la cooperación social (nuestra *racionalidad* y nuestra *razonabilidad).* Suponiendo que esto sea así, la pregunta inmediata es: ¿qué decisiones adoptarían los individuos una vez que se encontraran sometidos a tales condiciones?

Maximin. Rawls afirma que lo racional en esas condiciones sería adoptar una estrategia llamada *maximin* (una palabra que proviene de la jerga económica y que abrevia la expresión «maximizar el mínimo»). Recordemos que los principios y normas que vamos a adoptar van a influir a lo largo de nuestras vidas. Recordemos también que no sabemos cuál es nuestra concepción del bien ni qué lugar ocupamos en la sociedad. Puestos en estas condiciones, si negociamos mal podemos lamentarlo mientras vivamos. Por ejemplo, podemos condenarnos a nosotros mismos a no estar nunca en condiciones de realizar nuestra propia concepción del bien.

Visto este grave peligro, argumenta Rawls, lo racional no sería tratar de maximizar los posibles beneficios que podamos obtener, sino minimizar los riesgos a los que podríamos enfrentarnos en el caso de que nos toque estar entre los miembros menos favorecidos de la sociedad. En otras palabras, lo que tenemos que hacer no es maximizar nuestras ganancias potenciales en el caso de que nos toque ganar, sino

minimizar los eventuales daños en el caso de que nos toque perder. Rawls resume esta idea diciendo que tenemos que diseñar las instituciones de base de la sociedad bajo el supuesto de que nuestro lugar será asignado por nuestro peor enemigo. Supongamos, por ejemplo, que tenemos que decidir si habrá esclavitud en la sociedad. No es imposible encontrar argumentos en favor de esta idea. Por ejemplo, puede sostenerse que un régimen esclavista va a tener consecuencias beneficiosas sobre la productividad de la economía. Pero el modo correcto de evaluar este argumento, dice Rawls, consiste en imaginar que, una vez levantado el *velo de ignorancia,* vamos a descubrirnos del lado de los esclavos. Y vistas las cosas desde esta perspectiva, parece claro que el argumento de la productividad no es demasiado atractivo, porque nuestra condición de esclavos no nos dejaría ninguna oportunidad de aprovecharnos de sus beneficios.

Suponiendo que esta argumentación es aceptable (es decir, asumiendo que las condiciones de la *posición original* reflejan adecuadamente nuestras circunstancias de elección, y que en esas condiciones es racional adoptar la estrategia *maximin)* ¿qué decisiones específicas tomarían los individuos?

Libertades. Lo primero que establecerían, dice Rawls, es *la prioridad de las libertades fundamentales*. Esta prioridad significa que nuestras libertades básicas no pueden ser sometidas a ningún cálculo de utilidad. Tal como muestra el ejemplo de la esclavitud, no es aceptable que sacrifiquemos

una o varias de nuestras libertades fundamentales para aumentar la eficiencia económica. Y tampoco es aceptable sacrificar una o varias de esas libertades en beneficio de la igualdad material. La igualdad es un valor importante, pero debe ser alcanzada por vías que sean compatibles con el respeto de las libertades fundamentales.

Esto no significa que el ejercicio de nuestras libertades fundamentales no admita limitaciones, pero sí implica que *una libertad fundamental sólo puede ser limitada en beneficio de otra libertad fundamental*. Podemos, por ejemplo, mandar a un individuo a prisión (es decir, restringir severamente su libertad de movimiento) para asegurar la integridad física o el derecho a la vida de los demás. Pero no podemos hacerlo con el propósito de aumentar la eficiencia económica, instalando, por ejemplo, un régimen de trabajos forzados. Del mismo modo, podemos aceptar ciertas restricciones al ejercicio de la libertad de prensa como parte de las garantías que definen el debido proceso (por ejemplo, podemos prohibir que se publique el nombre de personas sin antecedentes en las primeras etapas de la investigación judicial). En cambio, no podemos prohibirle a alguien que manifieste sus opiniones políticas con la excusa de mejorar la imagen del país en el exterior.

La prioridad de las libertades fundamentales en el orden institucional es un asunto extremadamente importante, pero es sólo uno de los que deben ser resueltos por los participantes en la *posición original*. También será necesario deci-

dir cómo se van a distribuir las oportunidades y los recursos materiales entre los miembros de la sociedad. ¿Deberemos dejar que este asunto se resuelva solo, por la vía del libre intercambio entre los individuos, o debemos asignar a las instituciones políticas alguna función redistributiva?

Para resolver esta cuestión, Rawls vuelve a ponerse en la perspectiva de quienes deben tomar decisiones bajo el *velo de ignorancia*. Esos individuos no saben quiénes son, pero sí saben que están interesados en ejercer sus capacidades morales fundamentales (la capacidad de desarrollar e intentar poner en práctica una concepción del bien, y la capacidad de desarrollar un sentido de la justicia). Si por alguna razón se encontraran en un contexto en el que esas capacidades no pudieran ser ejercidas, su propia condición de agentes morales estaría en peligro.

Amenazas. Ahora bien, como el *velo de ignorancia* no les impide tener conocimientos generales, los individuos saben que esos contextos amenazadores existen. Por ejemplo, la capacidad de actuar como un agente moral quedaría hipotecada en un contexto en el que los individuos carecieran de la libertad necesaria para vivir en función de sus propias ideas acerca de cómo vale la pena vivir (justamente por eso es esencial la prioridad de las libertades fundamentales). Y nuestra agencia moral también quedaría hipotecada en un contexto en el que los individuos tuvieran libertad, pero carecieran de los medios mínimos indispensables para intentar poner en práctica un plan de vida. Por

tanto, así como los participantes en la *posición original* tienen razones para acordar la prioridad de las libertades fundamentales, también tienen razones para acordar que las instituciones aseguren a cada uno el acceso a los medios indispensables para intentar poner en práctica sus ideas acerca de cómo vale la pena vivir.

Este razonamiento resulta persuasivo para muchos lectores, pero todavía quedan dos importantes problemas por resolver. En primer lugar, ¿cómo identificar aquellos medios que deben ser asegurados a cada miembro de la sociedad? Y en el caso de que hayamos resuelto ese problema, ¿cómo encontrar un criterio que nos permita decidir el modo en que esos medios deben ser distribuidos?

Bienes. Rawls intenta resolver la primera cuestión desarrollando su teoría de los *bienes primarios*. Dado que nuestro punto de partida es el *hecho del pluralismo* (es decir, la radical diversidad de doctrinas morales, metafísicas y religiosas), no podemos aspirar a identificar un catálogo de cosas que sean reconocidas como buenas por todos los miembros de la sociedad. El dinero va a ser considerado esencial por algunos y despreciable por otros, la vida familiar va a ser vista como el más importante de los bienes humanos por unos y como una simple cárcel por otros. Cualquier intento de construir un catálogo de cosas buenas en sí mismas está inevitablemente destinado al fracaso. En cambio, lo que sí podemos hacer es intentar construir un catálogo de bienes que sirvan como medios para la realización de cual-

quier programa de vida, es decir, bienes que todo individuo racional debería desear, independientemente de cuáles sean sus ideas acerca de cómo vale la pena vivir. Si consiguiéramos identificar este conjunto de bienes, habríamos identificado aquello que haría falta distribuir.

Rawls cree que esta tarea es realizable. A su juicio, efectivamente hay un conjunto de bienes «que tendrán normalmente un uso, cualquiera que sea el plan racional de vida de una persona»[3]. Su argumento puede reconstruirse así: dado un individuo que ha conseguido desarrollar un plan de vida en función de sus ideas acerca de cómo vale la pena vivir, ese individuo tendrá un interés fundamental en poder realizarlo[4]; y dado el interés de ese individuo en realizar su plan de vida, entonces también tendrá un interés racional en acceder a ciertos «medios para todo propósito», esto es, a un conjunto de bienes cuya posesión nos pone en condiciones de intentar realizar cualquier plan. (Naturalmente, el acceso a esos *bienes primarios* no es en sí mismo una garantía de que podré realizar con éxito el plan de vida que he elegido. Por ejemplo: no llegaré a ser un gran pianista si no tengo acceso a un piano ni dispongo de tiempo para ejercitarme, pero puedo tener todo eso y no llegar a ser un gran pianista a causa de mis propias limitaciones).

---

3 En las notas siguientes se cita en primer lugar el número de página de la edición inglesa de Harvard University Press y luego el número de página de la edición castellana del FCE.

4 Rawls llama "el bien" de una persona a la realización exitosa de su plan de vida racional (Rawls, 1971, pág. 92; esp., pág. 114).

Catálogo. Rawls propuso en *Teoría de la justicia* un catálogo de estos *bienes primarios* que pueden ser vistos como medios para todo propósito (posteriormente el catálogo se modificó levemente). La lista incluye:

1. Los derechos y libertades fundamentales que protegen nuestra capacidad de elegir el tipo de vida que queremos vivir.
2. Los poderes y prerrogativas asociados a las diferentes funciones y posiciones que podemos ocupar en la sociedad (es decir, la capacidad de influencia y de decisión que podemos adquirir cuando nos situamos en ciertas posiciones).
3. La riqueza y el ingreso.
4. Algo relativamente abstracto que Rawls llama «las bases sociales del respeto de sí mismo», es decir, las condiciones que hacen posible «el sentimiento que una persona tiene de su propio valor, su firme convicción de que su concepción del bien, su proyecto de vida, vale la pena de ser llevado a cabo», junto a su «confianza en su propia capacidad, en la medida en que ello dependa de su propio poder, de realizar las propias intenciones». Las condiciones sociales que hacen posible este autorrespeto incluyen, por ejemplo, la ausencia de prácticas discriminatorias en materia de estudios, residencia o empleo.

La idea central de la teoría de los *bienes primarios* es que, cualquiera que sea el programa de vida que uno elija, tener más de estos bienes será siempre preferible a tener menos. Es perfectamente posible que ciertas personas hagan un mayor uso de algunos de estos bienes que de otros (por ejemplo, es probable que muchos no se interesen en acumular capacidad de decisión), pero lo importante es que todos puedan acceder a ellos, dentro de los límites que aseguren la misma posibilidad de uso por parte de los demás.

Distribución. Suponiendo que el primer problema ha quedado resuelto, queda por saber cómo hay que distribuir esos bienes. ¿Hay que hacerlo de manera estrictamente igualitaria de modo que todos tengan una cuota idéntica? ¿Hay que hacer una distribución inicial igualitaria y luego respetar las desigualdades que resulten del libre intercambio? ¿O las instituciones deben estar dispuestas a intervenir constante o periódicamente con el fin de asegurar cierta pauta de distribución?

Rawls se inclina por la tercera opción. Al discutir bajo el *velo de ignorancia,* los individuos saben que puede haber desigualdades importantes en la distribución de bienes primarios, y saben que a cada uno de ellos le puede tocar el lugar de quien ha recibido menos. También saben que si la disponibilidad de bienes primarios es demasiado exigua la posibilidad de ejercer su condición de agentes morales puede verse hipotecada. Todo esto hace que sea racional intentar maximizar la cuota mínima de bienes primarios que puede recibir un individuo (la célebre estrategia *maximin).*

Los individuos saben, sin embargo, que la distribución que maximiza el mínimo no es necesariamente la distribución igualitaria. Como vimos al principio, puede ocurrir que cierto grado de desigualdad resulte beneficiosa para todos los miembros de la sociedad, incluyendo aquellos que tienen una menor dotación relativa. Por tanto, *lo racional para los participantes en la posición original no es rechazar toda forma de desigualdad, sino aceptar solamente aquellas desigualdades económicas y sociales cuya existencia mejore la situación de los menos favorecidos.*

Lo que Rawls propone es adoptar un criterio que nos permita distinguir entre aquellas formas de desigualdad que vamos a reconocer como justas y aquellas que vamos a rechazar por injustas. Y el modo en que opera este criterio consiste en ligar el aumento de la desigualdad al mejoramiento de la suerte de los que están peor. Volvamos a nuestro modelo simplificado de sociedad integrada por dos individuos. Supongamos que nuestro punto de partida es una situación en la que el individuo A tiene 3 unidades y el individuo B 4, lo que arroja una suma total de 7:

A	3
B	4
	7

Supongamos ahora que es posible pasar a otra situación en la que los bienes se distribuyan del siguiente modo:

A	3
B	7
	10

En esta nueva situación la sociedad se ha vuelto más eficiente: la suma total pasó de 7 a 10. Sin embargo, sólo el individuo B se benefició con el cambio (su dotación personal pasó de 4 a 7). No sólo ocurre que la sociedad se volvió menos igualitaria, sino que este cambio no benefició al individuo A. Los individuos en la *posición original*, dice Rawls, no tendrán ninguna razón para aceptar esta nueva distribución, porque el punto de vista desde el que evaluarán las cosas es la posibilidad de que a cada uno de ellos le toque estar en el lugar de A.

Pero imaginemos ahora que existe una alternativa a la situación inicial que consiste en lo siguiente:

A	5
B	9
	14

En este caso, la sociedad se ha vuelto más eficiente (el producto creció a 14) y también se ha vuelto menos igualitaria. Pero el cambio benefició al individuo A, que antes tenía 3 unidades y ahora tiene 5. El individuo A sigue siendo el que está peor, pero su situación ha mejorado. Rawls argumenta que sería racional para los individuos en la *posición original* aceptar este cambio.

Rawls no identifica la justicia con la igualdad estricta. A su juicio, la injusticia no consiste en que haya desigualdades, sino en que haya «desigualdades que no benefician a todos». Y cada vez que se produzca una injusticia así entendida, las instituciones políticas tendrán una razón legítima para intervenir. Esto se debe a que, como dice la muy célebre primera línea de *Teoría de la justicia:* «La justicia es la primera virtud de las instituciones sociales, como la verdad lo es de los sistemas de pensamiento.»

Los principios de justicia

Toda la argumentación rawlsiana confluye en sus dos célebres principios de justicia, que han estado en el centro de la discusión filosófica durante los últimos treinta años. En general se los conoce como el *principio de igual libertad* y el *principio de diferencia,* aunque en realidad no se trata de dos principios sino de tres (el segundo combina el principio de diferencia propiamente dicho con un principio de igualdad de oportunidades). En el transcurso de los años, Rawls ha enunciado estos principios de varias maneras. La que se presenta aquí es la última formulación que aparece en *Teoría de la justicia*. Esta formulación será modificada en los años posteriores[5], pero lo que importa aquí es el estado de su pensamiento hacia 1971:

---

5 Esta enunciación se mantuvo intacta durante algo más de diez años. La última vez que aparece es en Rawls, 1982, págs. 161-62. La nueva formulación aparece por primera vez en «The Basic Liberties and their Priority», publicado en 1983 y luego incorporado como octavo capítulo de *Political Liberalism*. Para una justificación de los cambios, véase Rawls,1993, págs. 291 y 331.

(I) Cada persona tiene igual derecho al más amplio esquema de libertades fundamentales que sea compatible con un esquema similar de libertades para todos.
(II) Las desigualdades económicas y sociales han de satisfacer dos condiciones: tienen que (a) ser para el mayor beneficio de los miembros menos favorecidos de la sociedad; y (b) tienen que estar asociadas a cargos y posiciones accesibles a todos en condiciones de igualdad de oportunidades.

Estos principios se complementan con lo que Rawls llama dos reglas de prioridad. La primera de ellas (la *regla de prioridad de la libertad)* establece la preeminencia del principio I sobre el principio II. Esto significa que en ningún caso las libertades pueden ser sacrificadas para favorecer una distribución más justa. Rawls está a favor de las políticas redistributivas, pero bajo la exigencia innegociable de distribuir en libertad. La segunda regla (la *regla de prioridad de la justicia sobre el bienestar)* establece la preeminencia del principio IIb sobre el IIa. Podemos aceptar desigualdades asociadas al ejercicio de ciertas responsabilidades sociales (en el supuesto de que esas desigualdades van a beneficiar a los más desfavorecidos) pero sólo si previamente nos hemos asegurado de que el acceso a esos puestos de responsabilidad está abierto a todos los miembros de la sociedad en un marco de igualdad de oportunidades.

Los menos favorecidos. Para terminar de entender el significado de estos principios es importante aclarar el concep-

to de «miembros menos favorecidos de la sociedad». Lo esencial es percibir que se trata de un concepto relativo: los miembros menos favorecidos son aquellos que, *en un momento dado,* están en una situación desventajosa *respecto de los demás*. Para ilustrar esta idea volvamos a nuestra sociedad compuesta por dos individuos, donde el individuo A tiene 3 y el individuo B tiene 4 unidades de utilidad, y supongamos que alguien nos propone pasar de la situación I a la situación II:

	I	II
A	3	5
B	4	2
	7	7

Está claro que las dos situaciones son iguales en términos de eficiencia (en ambas el producto es 7) y que, si pasamos de una a otra, el individuo menos favorecido en la primera (es decir, el individuo A) mejora su posición. Pero el coste es que el individuo B está peor en la situación II de lo que estaba el individuo A en la situación I. El segundo principio de justicia propuesto por Rawls exige que, de estos dos posibles estados del mundo, elijamos el primero. Esto se debe a que el menos favorecido en la primera situación está mejor que el menos favorecido en la segunda, *aunque se trate de individuos diferentes*. Dicho en otras palabras: los individuos no pueden ser sacrificados en nombre de una justicia compensatoria.

La noción de *equilibrio reflexivo*

Rawls supone que, si ponemos a un conjunto de individuos racionales bajo las condiciones de la *posición original,* esos individuos van a llegar a un acuerdo en torno a estos principios de justicia y van a organizar sobre ellos la cooperación social[6].

Pero al mismo tiempo afirma que este procedimiento (al que llama «constructivo») no es suficiente para justificar las instituciones comunes. Además de realizar el esfuerzo constructivo mediante la ficción de la posición original, tenemos que someter las conclusiones a las que lleguemos a una dinámica muy particular que recibe el nombre de *equilibrio reflexivo.*

¿Por qué Rawls complica de este modo su esfuerzo de elaboración teórica? Porque es un autor de finales del siglo XX y, como tal, es consciente de la peligrosidad que encierra la construcción de sistemas fundados en grandes principios abstractos. Varias filosofías e ideologías que lograron formulaciones razonablemente coherentes terminaron conduciendo a terribles calamidades humanas. Por eso tenemos que actuar con prudencia a la hora de formular principios que orienten nuestra acción. Tenemos que estas seguros de que esos principios, además de ser inteligibles y coheren-

---

6 En realidad Rawls propone un procedimiento más complejo, a lo largo del cual las instituciones van quedando instaladas a medida que levantamos el velo de ignorancia. Véase el párrafo 31 de *Teoría de la Justicia.*

tes, no nos lleven a convertirnos en fanáticos políticos o morales.

Lo que Rawls necesita, entonces, es alguna instancia de control contra la cual contrastar los principios generales a los que llegó. Esto es algo que hicieron muchos filósofos del pasado cuando se preguntaban si las conclusiones a las que habían llegado eran compatibles con la voluntad de Dios, con la naturaleza humana o con el desarrollo de las fuerzas de producción. Pero, dado que todas esas instancias de control se han vuelto controvertidas en nuestras sociedades, ¿contra qué realidad podremos contrastar nuestras conclusiones?

Juicios bien ponderados

La respuesta de Rawls es: con nuestros *juicios morales bien ponderados*. Esta es una de las ideas básicas de su método de trabajo, de modo que vale la pena detenerse en ella.

Un primer rasgo de los *juicios morales bien ponderados* es que son juicios normativos particulares. No se trata de razonamientos morales complejos sino de reacciones puntuales de aceptación o rechazo. Formulamos esta clase de juicios cuando decimos cosas como: «esto es justo», «esto es injusto», «esto es aceptable» o «esto es inaceptable». Rawls llama «intuiciones morales» a esta clase de juicios. Pero la palabra «intuición» no se usa aquí en el sentido psicológico, sino en un sentido filosófico.

Un segundo rasgo de los *juicios bien ponderados* es que son juicios particularmente confiables. Se trata de juicios

«emitidos bajo condiciones favorables para el ejercicio del sentido de la justicia y, por tanto, en circunstancias en las cuales no son de aplicación las excusas y explicaciones más comunes. Se presume entonces que la persona que formula el juicio tiene la capacidad, la oportunidad y el deseo de llegar a una decisión correcta...»[7]. Esto no significa que sean infalibles, pero sí significa que nos ofrecen el mayor grado de confiabilidad al que podemos aspirar los seres humanos. Estamos seguros de que no hay mayores ausencias de información, que hemos examinado todas las caras del problema y que lo hemos hecho con el deseo sincero de llegar a una conclusión correcta.

Sentido de justicia. Una vez que aceptamos que los hombres no tenemos línea directa con un mundo de verdades eternas, los *juicios morales bien ponderados* se vuelven hechos particularmente relevantes de nuestro universo moral: en ellos se deposita un «sentido de justicia» compartido por los individuos racionales y razonables que intentan encontrar principios justos para ordenar la convivencia. No se trata de juicios trascendentales en el sentido kantiano, pero es lo más que podemos acercarnos a algo semejante. Por tanto, si un principio general que hemos formulado entra en colisión con nuestros *juicios bien ponderados,* eso significa que tenemos problemas. O bien nuestro principio debe ser

7 Aquí, como en otros párrafos, utilizo mi propia traducción.

modificado, o bien debe serlo nuestro juicio o bien deben serlo ambos.

Justamente en esto consiste la mecánica del *equilibrio reflexivo*. Los principios y normas a los que lleguemos por la vía de razonar en las condiciones de la *posición original* deben ser cotejados con nuestros *juicios morales bien ponderados*. Si no hay colisión, esto nos permite seguir avanzando en la elaboración de nuestra teoría moral. Si hay colisión, o bien deberemos modificar nuestros principios generales o bien deberemos modificar nuestros juicios. Esto último es lo que ocurre cuando los principios generales pueden explicar el conflicto y ofrecer un juicio alternativo.

Vaivén. La elaboración de una teoría moral consiste así en un movimiento de vaivén entre nuestros *juicios morales bien ponderados* y los principios generales que vamos formulando[8]. El *equilibrio reflexivo* se alcanza cuando los principios y las intuiciones están (al menos provisionalmente) «en línea», es decir, cuando no hay disonancias importantes entre los principios que elegimos y los juicios que han sobrevivido a este proceso de confrontación.

Como es fácil imaginar, la aplicación detallada de este método plantea problemas complejos. Por ejemplo, ¿cuál es el grado mínimo de coincidencia con nuestros *juicios bien*

8 Rawls distingue entre un equilibrio reflexivo estrecho y un equilibrio reflexivo amplio. Sobre el punto véase Rawls, 1971, págs. 49-50 (esp., pág. 69) y Van Parijs, 1991, pág. 73 y sigs.

ponderados que debemos exigirle a una teoría? ¿Qué condiciones debe cumplir un *juicio bien ponderado* para que pueda ser considerado compartido? O para plantear una cuestión todavía más difícil: si partimos de un mismo conjunto de *juicios morales bien ponderados*, ¿la práctica correcta del equilibrio reflexivo debería conducirnos a elegir los mismos principios?[9]. Estos problemas son metodológicamente importantes, pero no es necesario discutirlos para entender el sentido de la propuesta rawlsiana. Lo que me importa señalar aquí es que, para aceptar en términos generales la validez de su método, es necesario haber aceptado dos supuestos.

El primer supuesto se remonta a Aristóteles y consiste en observar que nuestra vida moral no consiste en deducir juicios y decisiones particulares a partir de un puñado de principios generales[10]. Simplemente, las cosas no ocurren de ese modo. A veces estamos seguros de no querer abandonar un principio general que guía nuestra acción, y a veces estamos dispuestos a revisar un principio general que hemos aceptado porque nos conduce a conclusiones particulares que consideramos inaceptables. En palabras de Rawls, «es un error pensar que las concepciones abstractas y los principios generales están siempre por encima de nuestros juicios particulares. Estas dos caras de nuestro

---

9 *Véase* 1971, págs. 49-50, esp., págs. 69-70.

10 Rawls reconoce el parentesco con Aristóteles en 1971, pág. 51n, esp., pág. 71n.

pensamiento práctico [...] son complementarias». La conclusión que extrae Rawls es que la función que debe desempeñar la filosofía moral consiste en organizar y refinar un sentido de la justicia que está en parte estructurado a partir de principios generales y en parte estructurado en torno a juicios particulares que nos parecen especialmente bien fundados.

El segundo supuesto de esta propuesta metodológica es que, cuando nos ocupamos de esa parte de la teoría moral que llamamos filosofía política, nuestro lugar no está fuera del cuerpo de ciudadanos sino dentro de él. El *equilibrio reflexivo* es viable porque al hacer filosofía asumimos como nuestros los principios e intuiciones que están implícitos en la cultura política compartida. El filósofo político no es un observador externo que se limita a anotar los errores y carencias de los ciudadanos, ni un crítico que se encarga de denunciar las maquinaciones de las que son víctimas. Es simplemente un ciudadano tan leal como el que más a los principios de la democracia liberal, que intenta razonar en diálogo con los demás ciudadanos.

El esfuerzo de Rawls por pensar en las condiciones de la *posición original* no debe ser visto como un intento de pensar desde cero un nuevo tipo de sociedad, sino como un intento de depurar y extraer lo mejor de nuestras tradiciones cívicas y morales. Este es un punto importante porque pone de relieve una de las características más notables: Rawls

escribió su primer libro en un momento de hondo descrédito de la democracia liberal (recordemos que *Teoría de la justicia* apareció en 1971), pero su pensamiento siempre se mantuvo fiel a ella. Una de sus intuiciones centrales es que las instituciones de la democracia liberal son una construcción moralmente admirable. El fin para el que fueron hechas no es asegurar el orden a cualquier precio ni oponerse al cambio, sino hacer posible una convivencia social justa en el respeto de la diversidad. Lo que tenemos que hacer no es destruirlas, sino recuperar su impulso original. En medio de las múltiples crisis de fines de los años sesenta, Rawls se proponía hacer una defensa no conservadora de la democracia liberal.

CAPÍTULO 2

Tres décadas de discusión

La publicación de *Teoría de la justicia* fue seguida de una larga serie de debates que en parte se sucedieron y en parte se superpusieron en el tiempo. El primero de ellos se planteó entre Rawls (junto a otros filósofos que defendían puntos de vista similares) y los defensores del utilitarismo. El segundo se estableció entre Rawls (acompañado de los demás liberales igualitaristas) y los liberales libertaristas. El tercero se planteó entre Rawls (acompañado de otros igualitaristas moderados) y los igualitaristas radicales. El cuarto y último se entabló entre Rawls (junto a otros liberales más o menos individualistas) y los filósofos comunitaristas. Describir brevemente cada una de estas confrontaciones puede ayudarnos a percibir la influencia que Rawls fue ganando con el paso del tiempo, así como a entender por qué esa influencia terminó por sobrepasar los límites del mundo anglosajón.

El debate con los utilitaristas

El debate con el utilitarismo fue la única de las discusiones iniciadas por el propio Rawls. De hecho, *Teoría de la justicia* puede ser vista como el intento de construir una alternativa teórica a esta forma de pensamiento. La importancia que tuvo esta discusión no es fácil de percibir para los lectores hispanohablantes, porque el utilitarismo nunca tuvo demasiada aceptación en Europa continental ni en América Latina. Pero las cosas eran muy distintas en el mundo anglófono, donde el utilitarismo llevaba acumulados dos siglos de predominio en el pensamiento moral.

La familia utilitarista. No es sencillo resumir lo que los utilitaristas tienen en común. Se trata de una familia grande y heterogénea que incluye a figuras clásicas como Jeremy Bentham (1748-1832), James Mill (1773-1836), su hijo John Stuart Mill (1806-1873) y Henry Sidgwick (1838-1900), junto a autores contemporáneos como el británico Richard M. Hare y el húngaro John Harsanyi. Pero puede afirmarse, al menos de manera general, que en el pensamiento de todos ellos aparecen dos rasgos distintivos.

En primer lugar, los utilitaristas identifican al bien común con la suma (o el promedio) del bienestar alcanzado por los individuos. A la hora de identificar lo que beneficia o perjudica a la sociedad, sólo debemos fijarnos en los efectos que una decisión puede tener sobre el bienestar de cada persona, una vez que se ha aceptado que cada miem-

bro de la sociedad cuenta como uno y no más que uno. Ideas como las de interés nacional, voluntad general o bien supremo son rechazadas por poco confiables y oscuras.

En segundo lugar, a la hora de identificar lo que vamos a considerar como el bien de cada uno, lo único que cuenta es lo que cada persona identifica como su propio bien. El utilitarismo asume que cada individuo debe ocuparse de identificar lo que mejor le conviene, no porque los individuos no podamos equivocarnos en este punto sino porque cualquier alternativa es peor[11]. En el mundo utilitarista no hay lugar para reyes filósofos ni tutores benevolentes.

A partir de esta doble base, el utilitarismo propone que las decisiones sociales se apoyen en un cálculo de las pérdidas y ganancias de bienestar experimentadas por los individuos. La decisión a adoptar será aquella que maximice el bienestar total o el bienestar promedio del grupo sobre el que vaya a ser aplicada. El objetivo de un buen gobierno consiste en buscar —según una célebre fórmula del escocés Francis Hutcheson— *la mayor felicidad para el mayor número.*

Claridad y tecnificación. Esta manera de abordar las cuestiones normativas tiene dos méritos importantes. En primer lugar, tiene el atractivo de la claridad: frente a las imprecisiones y oscuridades de otras teorías morales, el utilitarismo propone un criterio general de decisión. En se-

11 El utilitarismo de Stuart Mill no exhibe este segundo rasgo, lo que lo convierte en una de las versiones más problemáticas de la doctrina.

gundo lugar, este abordaje permite una «tecnificación» de las decisiones morales y políticas: una vez que se ha adoptado el criterio de maximizar el bienestar es fácil servirse de instrumentos matemáticos para determinar con relativa precisión cuál es la opción a privilegiar. Esto explica por qué el utilitarismo ha tenido tanta influencia sobre la teoría y sobre la práctica económicas.

Dificultades. Pero este atractivo inicial no impide que aparezcan dificultades. Algunas de ellas son de carácter técnico. Por ejemplo: si asumimos que el bienestar debe ser maximizado, ¿cómo vamos a medir las variaciones de bienestar? ¿Cómo vamos a hacer comparaciones interpersonales? ¿Cómo vamos a comparar intensidades en el caso de que sea necesario hacerlo? Estas preguntas no son fáciles de responder. Supongamos que mi deseo de dormir la siesta entre en colisión con el deseo de mi vecino de escuchar rock a todo volumen. ¿Qué decisión deberíamos tomar para maximizar la utilidad total? Si pudiéramos asignarle puntos a la satisfacción de cada preferencia en función de una única escala, podríamos calcular cuál es la combinación que diera una mayor puntuación total. Pero el problema es que no tiene demasiado sentido decir que la satisfacción de mi deseo de dormir la siesta está tantos puntos por encima o por debajo del deseo de mi vecino de escuchar rock, como si se tratara de dos cantidades de la misma cosa.

Estas dificultades obligaron a los utilitaristas a sofisticar su doctrina. De este modo, el utilitarismo hedonista al es-

tilo Bentham, preocupado por los sentimientos de placer y dolor, dejó lugar a un utilitarismo de las preferencias que sólo atiende a aquello que los individuos dicen desear cuando se enfrentan a alternativas excluyentes. Y este utilitarismo de las preferencias fue a su vez desafiado por una versión más sofisticada (el utilitarismo de las preferencias racionales), según la cual el bienestar no consiste en la satisfacción de cualquier tipo de preferencia, sino de aquellas basadas en información correcta y que no estén afectadas por vicios de razonamiento[12].

Esta flexibilidad teórica para responder a las dificultades permitió resolver algunos problemas pero no hizo amainar las críticas. Esto se debe a que las objeciones al utilitarismo no sólo son de carácter técnico sino también normativo. Quienes lo atacan en este plano no dicen que sea una doctrina inaplicable sino que, aun en el caso de que fuera aplicable, conduciría a resultados indeseables. Y este era justamente el punto de vista de Rawls cuando redactaba *Teoría de la justicia*.

Críticas al utilitarismo. Rawls las resume en dos puntos fundamentales: le acusa de conducir, al menos potencialmente, a desigualdades extremas y de no ofrecer suficientes garantías de respeto a los derechos individuales. Ambas críticas se resumen en un lema que Rawls repitió incansa-

---

12 Para una discusión más detallada, véase Brandt, 1979; Sen & Williams, 1982; Van Parijs, 1991; Goodin, 1991; Audard, 1996 y 1999.

blemente: el utilitarismo no respeta la separación entre las personas, es decir, no tiene en cuenta la exigencia kantiana de ver a los demás como fines en sí mismos y no como medios para alcanzar nuestros propios fines.

La primera parte de la acusación ha sido suscrita por un gran número de autores. En esencia afirma que, al atender exclusivamente a la suma total o al promedio de bienestar, el utilitarismo no se preocupa del modo en que el bienestar es distribuido dentro de la sociedad. Para entender mejor esta afirmación, volvamos a considerar nuestro modelo simplificado de sociedad y asumamos que debemos elegir entre dos estados del mundo. En el primer estado, el individuo A tiene 3 unidades de utilidad y el individuo B tiene 4. El producto total es 7. En el segundo estado del mundo, el individuo A sólo tiene 2 unidades y el individuo B tiene 19, por lo que el producto total se triplica.

	I	II
A	3	2
B	4	19
	7	21

¿Cuál de las dos situaciones debemos elegir? Si somos utilitaristas consecuentes tenemos que optar por la II, porque tanto la suma global como el promedio son mayores que en la I. No importa que la situación I sea más igualitaria que la situación II, ni que la situación de A empeore al pasar de I

a II, ni que el individuo B se apropie de todo el fruto del crecimiento (y aún más). Lo único que cuenta es que la situación II permite maximizar la utilidad.

La segunda objeción de Rawls consiste en decir que el utilitarismo puede terminar justificando graves violaciones de los derechos individuales. Si lo que cuenta es la utilidad total o promedio, la satisfacción o el sufrimiento de cada individuo no tiene otro valor que lo que agrega o quita al conjunto. Y esto supone el riesgo de instrumentalizar a una parte de la sociedad. Por ejemplo, este razonamiento permitiría justificar que esclavizáramos al 5 por 100 de la población si pudiéramos demostrar que habría una mejora en la competitividad de la economía, o que condenemos de vez en cuando a un inocente para reforzar la imagen de que nada escapa al brazo de la justicia.

El modo en que Rawls define las características de los participantes en la *posición original* está claramente orientado a evitar estos riesgos. Como los individuos están interesados en el ejercicio de su agencia moral no aceptarán ninguna propuesta que someta el respeto de las libertades fundamentales a un cálculo de utilidad. Y como saben que el ejercicio de esas capacidades requiere el acceso a un conjunto básico de oportunidades y de medios materiales, no podrán ver a la maximización de la eficiencia como un bien en sí mismo, sino como un objetivo que sólo adquiere valor en la medida en que se lo asocie a ciertas exigencias distributivas. Toda su teoría tenía, pues, un claro tinte antiutilitarista.

Respuestas utilitaristas. Los autores utilitaristas se dieron perfecta cuenta de lo que estaba ocurriendo e intentaron una doble reacción. Por distintos caminos intentaron mostrar que: (a) los ataques contenidos en *Teoría de la justicia* eran infundados, y (b) la teoría propuesta por Rawls era incapaz de sostenerse por sí misma, dado que los participantes en la posición original no elegirían los principios rawlsianos sino, precisamente, los principios utilitaristas.

La discusión que se desató alcanzó altos niveles de tecnicidad, pero algunos ejemplos concretos permiten hacerse una idea del modo en que se procesó. Consideremos, por ejemplo, la acusación de que el utilitarismo podría conducir a formas extremas de desigualdad. Este temor es infundado, replicaron algunos utilitaristas, porque no tiene en cuenta el principio de utilidad marginal decreciente, es decir, aquel principio que afirma que cuanto más posee una persona de algo, menos bienestar le añade una unidad suplementaria de lo que ya posee.

Imaginemos dos individuos que ganan respectivamente 600 y 10.000 dólares al mes, y supongamos que debemos decidir a quién le damos 500 dólares suplementarios. Para el que gana 10.000, esos 500 dólares suplementarios sólo representan un 5 por 100 de sus ingresos. Para el que gana 600, esos 500 dólares suponen casi una duplicación de sus ingreso al mes. Es probable que la primera persona use ese dinero para darse algún gusto suntuario, en tanto que la segunda lo emplee para satisfacer una necesidad impor-

tante. Por tanto, si la utilidad marginal decrece cuanto más tiene una persona (es decir, si la cantidad de bienestar que se agrega tiende a ser cada vez más pequeña), el utilitarismo debería impulsar aquellas redistribuciones que favorezcan a los que tienen menos. La acusación de conducir a una creciente desigualdad sería entonces infundada.

Las preferencias adaptativas. El argumento tiene su fuerza, pero Rawls y los demás críticos del utilitarismo no lo consideran suficiente. Para ellos sigue siendo cierto que el utilitarismo conduce a una mayor desigualdad a causa de un conjunto de mecanismos que tienden a neutralizar el efecto de la utilidad marginal decreciente. Entre estos mecanismos destaca el de las *preferencias adaptativas,* que fue especialmente trabajado por Jon Elster. Veamos brevemente cómo funciona.

Uno de los axiomas del utilitarismo dice que no debemos atender a lo que la gente debe preferir sino a lo que efectivamente prefiere. Ahora bien, lo que cada uno de nosotros prefiere está fuertemente condicionado por nuestras experiencias anteriores y por las posibilidades reales de alcanzarlo. Cuando una persona pobre piensa en lo que le gustaría comer no piensa en lo que le gustaría comer si tuviera los medios, la experiencia previa y el gusto refinado de un *gourmet,* sino en lo que le gustaría comer dados sus medios, su experiencia previa y sus propios gustos. Esto hace que una persona pobre pueda extraer un gran placer de un plato que sólo provocaría malestar en un consumidor

sofisticado. Del mismo modo, lo que la gente considera una vivienda confortable, un empleo atractivo, una buena educación o un salario digno varía enormemente según sus experiencias previas y el lugar que ocupa en la sociedad.

Si esto es así, entonces es cierto que el utilitarismo podría terminar convalidando formas extremas de desigualdad. De acuerdo a su propia lógica, la mejor manera de maximizar el bienestar consistiría en darle comida barata al pobre y comida cara al rico, casas de mala calidad a los desfavorecidos y de buena calidad a los acomodados, educación de baja calidad a los incultos y de alta calidad a los cultivados. La utilidad marginal decreciente no es suficiente para compensar esta dificultad.

Medios y afines. La misma clase de discusión se desarrolló en torno a la acusación que sostiene que el utilitarismo podría llevarnos a instrumentalizar a una parte de la sociedad. Algunos utilitaristas intentaron resolver este problema por la vía de recuperar una vieja distinción entre un *utilitarismo del acto* y un *utilitarismo de la regla*. Si consideramos el modo en que cada acción particular influye sobre la utilidad global, entonces puede ocurrir que terminemos utilizando a una parte de nuestros conciudadanos como simples medios para nuestros fines. Pero el riesgo desaparece si en lugar de preferir aquellas acciones particulares que maximizan la utilidad global preferimos aquellas acciones que, convertidas en reglas generales, tienen el efecto de maximizar la utilidad global.

Ejemplo. Este trabalenguas puede explicarse de un modo sencillo. Supongamos que tengo una deuda con una persona inmensamente rica. Pagar esa deuda me supondría un enorme sacrificio, mientras que a mi acreedor le resulta relativamente indiferente recuperar su dinero o no. En estas condiciones parece claro que *no* pagar la deuda es la acción que maximiza la utilidad global: el aumento de mi bienestar al verme librado de esa carga va a ser mayor que la eventual disminución del bienestar de mi acreedor por no poder cobrar. Ahora bien, si cada individuo hiciera este tipo de razonamiento el resultado sería una destrucción de los sistemas de crédito. Nadie tendría razones para prestar dinero (porque no habría suficientes garantías de pago) ni nadie que estuviera necesitado de dinero tendría posibilidades de conseguirlo. Por tanto, la regla que exige pagar *siempre* las deudas va a ser a largo plazo más beneficiosa en términos de bienestar que la regla que exige pagarlas sólo cuando ese acto aumenta la utilidad global en un momento dado.

Ciertos utilitaristas creen que esta distinción bastaría para asegurar la protección de los derechos fundamentales de todos individuos, ya que la utilidad global disminuiría si cada uno de nosotros corriera permanentemente el riesgo de ser instrumentalizado (por ejemplo, si todos estuviéramos permanentemente expuestos a ser encarcelados por crímenes que no cometimos). La maximización de la utilidad global en el largo plazo requeriría la protección de los

derechos. Pero, una vez más, esta respuesta no ha dejado satisfechos a los críticos del utilitarismo[13]. En una sociedad donde exista una pequeña minoría racial se podría aprobar una regla que dijera que nadie puede ser obligado a trabajar gratis para otro individuo, a menos que el obligado pertenezca a esa minoría. Y no es imposible que esta norma, aplicada sistemáticamente, condujera a un aumento de la utilidad global, al menos en el caso de que la mayoría no viera ofendida su sensibilidad moral a causa del tratamiento que reciben los miembros minoritarios.

Ofensiva utilitarista. Las discusiones de este tipo se han vuelto muy complejas y se prolongan hasta hoy[14]. Pero esta fue sólo una de las dimensiones en las que se desarrolló el debate, porque los utilitaristas no se limitaron a defenderse sino que pasaron a la ofensiva. Su argumento principal consistió en decir que la teoría propuesta por Rawls era indefendible en términos racionales, ya que los individuos puestos en las condiciones de la *posición original* no elegirían los principios rawlsianos sino los principios utilitaristas. Esta crítica fue formulada en términos clásicos por John Harsanyi, un filósofo que consiguió el premio Nobel de Economía.

Rawls había sostenido que, al verse colocados en las con-

---

13 Ni a algunos utilitaristas destacados, como Richard Hare.

14 Para una descripción de etapas posteriores de la discusión véase, por ejemplo, Van Parijs, 1991, págs. 29 y sigs., y Audard, 1999.

diciones de la *posición original,* los individuos buscarían asegurar que el peor resultado fuera el menos malo posible, en lugar de optar por una estrategia que les prometiera mayores premios potenciales al coste de admitir cierto riesgo de sufrir pérdidas serias.

Riesgo o no riesgo. Esta es una estrategia extremadamente conservadora, sostiene Harsanyi, que no refleja el modo en que actuamos en la vida cotidiana ni el modo en que solemos valorar la racionalidad de los actores. En general estamos dispuestos a asumir riesgos relativamente importantes para obtener mayores beneficios. No tiene sentido tachar a una persona de irracional porque prefiera trabajar en el sector privado (donde tiene menos estabilidad laboral pero mayores expectativas salariales) en lugar de hacerlo en el sector público (donde tiene un empleo seguro pero bajas expectativas de ingresos). Esa persona es perfectamente racional y su propia racionalidad le ha llevado a desarrollar una estrategia que incluye cierta cuota de riesgo. Rawls, en cambio, identifica la racionalidad con una total aversión al riesgo, lo que es inaceptable.

Pero Rawls tiene argumentos para contrarrestar esta crítica. Asumir ciertos riesgos puede ser racional en la vida de todos los días, pero no lo es cuando se trata de definir las *instituciones de base* de la sociedad. Si en este caso nos alejamos de la estrategia *maximin,* lo que podemos perder no es un puesto de trabajo ni cierta cantidad de dinero, sino la posibilidad misma de vivir como agentes morales capaces

de ejercer nuestras *capacidades fundamentales*. Es perfectamente racional que yo corra riesgos en materia de empleo si con eso puedo mejorar mi nivel de ingresos, pero no es racional que hipoteque mi propia condición de agente moral para obtener ventajas de las que sólo podré disfrutar si puedo actuar como agente moral. Cuando es esto lo que está en juego, la aversión al riesgo debe ser absoluta.

El debate con los libertaristas

Un segundo frente de lucha al que muy pronto debió atender Rawls fue el abierto por los filósofos libertaristas. Estos autores son llamados con frecuencia «liberales de derecha», pero esta caracterización no permite apreciar plenamente su punto de vista.

Es verdad que, en muchos aspectos, los libertaristas coinciden con lo que en política suele llamarse «derecha». Por ejemplo, defienden vigorosamente la propiedad privada (incluido el derecho irrestricto de herencia), se oponen a casi toda forma de intervención estatal, tienen una confianza prácticamente ilimitada en el mercado y no vacilan en describir los impuestos como una forma de robo. Pero en muchos otros aspectos los libertaristas coinciden con lo que suele llamar «izquierda» y aun «izquierda radical». Por ejemplo, se oponen al servicio militar obligatorio, reclaman la eliminación de todos los obstáculos a las migraciones, defienden encarnizadamente la libertad de expresión (aunque se use para difundir pornografía) y rechazan toda forma de intromisión del

Estado en la vida privada de las personas (por ejemplo, en relación al consumo de drogas).

El resultado de esta doble condición es que todo intento de situarlos en un punto preciso del espectro derecha-izquierda genera serias dificultades.

Robert Nozick es, probablemente, el más célebre de los filósofos libertaristas estadounidense. Se trata del mismo colega de Harvard que saludó la aparición de *A Theory of Justice* diciendo que de allí en adelante los filósofos debían trabajar en ese marco o bien explicar por qué no lo hacían. Pero a pesar de toda la admiración que Nozick sentía por Rawls, estaba en desacuerdo con él en casi todo. Su libro más influyente es un intento de mostrar por qué Rawls se equivoca y en qué dirección debería orientarse la búsqueda de principios normativos.

Nozick empieza por observar que las teorías de la justicia pueden ser de dos tipos. Para algunas de ellas, un estado de cosas es justo o injusto según cómo quede conformada la matriz de distribuciones (es decir, según quién se quede con qué una vez que se han realizado las transferencias). Se trata del tipo de valoración que hacíamos al recurrir a nuestro modelo de sociedad ideal con dos individuos. Nozick llama a este grupo *teorías de estado final*. Lo único que cuenta para ellas es el estado de cosas al que se llega, independientemente del modo en que nos hayamos aproximado. Para un utilitarista lo que cuenta es si el estado al que se llega maximiza la utilidad global. Para Rawls, lo

único que cuenta es si el estado al que se llega mejora la situación de los menos favorecidos.

Pero otras teorías de la justicia funcionan de manera diferente. En estos casos, lo que determina si una situación es justa o injusta no es el estado final al que se llega sino cómo se llegó a él. Una situación es justa si se llegó a ella de un modo justo, y es injusta si ocurrió lo contrario. Nozick llama *teorías históricas,* o *retrospectivas,* a esta clase de teorías de la justicia. Es el tipo de teoría que aplicamos intuitivamente cuando decimos que es justo que una persona haya accedido a tal cosa porque se lo merece. Lo que cuenta aquí no es la comparación entre lo que posee esa persona y lo que poseen las demás, sino el modo en que esa persona llegó a tener lo que tiene.

El argumento crucial de Nozick consiste en decir que las teorías históricas o retrospectivas son preferibles a las teorías de estado final. Si esto es correcto, la teoría de Rawls habría recibido un duro golpe. Pero, ¿por qué deberíamos preferir las teorías históricas? La respuesta de Nozick es que son las únicas compatibles con el respeto de la libertad. Las teorías de estado final sólo pueden aplicarse si estamos dispuestos a coartar permanentemente la libre iniciativa de los individuos.

Chamberlain. Para defender esta idea, Nozick lanza uno de los desafíos más célebres de la historia de la filosofía política. Elijan, dice, la teoría de estado final que más les atraiga. Puede tratarse de una distribución que maximice la

utilidad, de una distribución que aplique el principio de «a cada quien según sus necesidades» o de una distribución que aplique el principio de diferencia de Rawls. Luego realicen la distribución de acuerdo con esa teoría, de manera que puedan obtener el estado de cosas que consideren preferible. «Ustedes eligen, nos desafía Nozick, y yo acepto esa distribución como punto de partida».

Ahora imaginemos que Wilt Chamberlain es un jugador de baloncesto que está realizando una excelente temporada deportiva. Supongamos que Chamberlain firma con su club un contrato por un año que establece que en cada partido que su equipo juegue como local, 25 centavos del precio de cada entrada vendida serán para él y el resto para el club. Supongamos también que el público está enormemente entusiasmado con el gran Wilt y que, al cabo del año, un millón de espectadores fueron a verlo cuando jugaba como local. Esto supone que Chamberlain ganó 250.000 dólares, una cifra muy superior al ingreso anual medio de sus conciudadanos. (Las cifras del ejemplo son insignificantes en la era de Michael Jordan, pero no debe olvidarse que el libro de Nozick fue publicado en 1974.)

Al cabo de la temporada Wilt se ha vuelto más rico que la mayoría de los estadounidenses, pero esto ocurrió como resultado de una serie de transferencias estrictamente voluntarias: nadie pagó su entrada sin querer hacerlo, nadie se sintió estafado, los espectadores estaban entusiasmados y pensaban que su dinero fue bien empleado. Podrían haber usado esos 25 centavos para otras muchas

cosas (ahorrar, invertir, comprar libros de filosofía), pero prefirieron dárselos a Wilt Chamberlain y están satisfechos de haberlo hecho.

Poseer y decidir. El resultado de este proceso es que la distribución que Nozick aceptó como punto de partida ha desaparecido. Y sin embargo, no se ve por qué deberíamos considerar la nueva situación como injusta. ¿En función de qué razones el Estado o cualquier otro agente podría retirar parte de lo que Wilt Chamberlain ha ganado para redistribuirlo entre los espectadores? Esta clase de confiscación sería necesaria si quisiéramos mantener el estado final que hemos elegido, pero eso nos obligaría a ir contra las decisiones libres de individuos que han hecho un uso racional y voluntario de los recursos de los que disponían. Y poseer algo, nos recuerda Nozick, es poder decidir qué hacemos con lo que poseemos.

En una sociedad compleja, argumenta Nozick, los casos como el de Wilt Chamberlain se multiplican. En consecuencia, las teorías de estado final sólo pueden aplicarse si estamos dispuestos a «intervenir continuamente para impedir que la persona transmita recursos como quiera, o a intervenir en forma continua (o periódica) para tomar recursos de algunas personas que otras, por alguna razón, decidieron trasmitirles a ellas». En otras palabras, las teorías de estado final sólo pueden aplicarse si estamos dispuestos a limitar severamente la libre iniciativa y a desnaturalizar el concepto de propiedad.

El razonamiento de Nozick es en realidad algo más complejo de lo sugerido hasta aquí. En su opinión, las teorías de la justicia no sólo pueden dividirse en históricas y de estado final, sino también en pautadas y no pautadas. Una teoría de la justicia es pautada si establece criterios específicos que determinan el modo en que debe variar la distribución. No lo es si se limita a establecer un principio (o un conjunto de principios) que especifican algunas condiciones que debe cumplir una distribución justa sin que se pueda derivar de ellos una distribución específica.

Tanto las teorías de estado final como las teorías históricas pueden ser pautadas o no pautadas. Una teoría que exija una mínima dotación de recursos para cada individuo es de estado final y no pautada (esa teoría permite identificar el conjunto de distribuciones que podrán ser consideradas justas, pero no nos da elementos para elegir una en particular). Una teoría que proponga aumentar la dotación de bienes o de recursos a medida que aumenta el cociente intelectual de los individuos es de estado final y pautada. El principio de «a cada uno según su trabajo» es histórico y pautado (tiene en cuenta el modo en que se llegó a la distribución actual, pero además establece una métrica que permite fijar la distribución). El principio «cada uno puede conservar lo que adquirió legítimamente» es histórico pero no pautado (es decir, no nos dice nada acerca de cómo quedarán finalmente distribuidos los bienes).

Nozick argumenta que no sólo las teorías de estado final, sino también las teorías históricas pautadas, se enfrentan a la dificultad que describe el ejemplo de Wilt Chamberlain. Tomemos el ejemplo de «a cada uno según su trabajo». Supongamos que realizamos una distribución en base a este principio y que luego ponemos a rodar la historia de Wilt Chamberlain. Perfectamente puede ocurrir que alguien que ha recibido cierta cantidad como resultado de su trabajo termine el año más pobre porque fue a ver jugar a Wilt todas las semanas. Chamberlain, por su parte, se habrá enriquecido por encima de lo que establece el criterio. La pauta sólo podrá mantenerse si intervenimos para corregir la situación.

La conclusión de Nozick es que una teoría de la justicia sólo será aceptable si es al mismo tiempo *histórica y no pautada*. Él mismo va a proponer una teoría de este tipo que lo llevará a alinearse con el Estado mínimo. La sociedad no tiene ningún derecho a disponer de los bienes que los individuos han adquirido mediante procedimientos legítimos. Por tanto, las únicas formas de redistribución tolerables son las que se realizan de forma voluntaria. Toda intervención que vaya más allá de este límite es violadora de los derechos individuales. En consecuencia, «el Estado mínimo es el Estado más extenso que se puede justificar». Se trata de un retorno a la idea del Estado juez y gendarme.

Pero lo que importa aquí no es la teoría de la justicia

que propone Nozick, sino los argumentos que ha empleado para descalificar la teoría rawlsiana. Si Nozick tiene razón, Rawls está en serias dificultades. Pero ¿es sustentable la argumentación de Nozick? Rawls y muchos liberales igualitaristas creen que no, pero tuvieron que hacer una amplia argumentación para justificar esta afirmación.

En opinión de los liberales igualitaristas, la teoría de Nozick sólo puede sostenerse en la medida en que se acepten dos supuestos implícitos. El primero es que los hombres tenemos derechos naturales, es decir, derechos que preexisten al Estado. Por ejemplo, mi derecho a caminar por donde me plazca existiría antes del contrato social y se traduciría en el reconocimiento por parte del Estado del derecho a la libre circulación. El segundo supuesto es que estos derechos anteriores al Estado son absolutos, es decir, son derechos que nadie puede limitar. Mi derecho a disponer de mi propio cuerpo implica que puedo hacer con él lo que quiera, incluyendo la posibilidad de venderlo o de alquilarlo. Para Nozick, tener un derecho significa tenerlo sin ninguna limitación.

Estos dos supuestos son necesarios para que la argumentación de Nozick funcione, pero el problema es que ambos son problemáticos. Empecemos por el primero. Nozick da por obvio que los hombres tenemos derechos que preexisten al Estado. La primera frase de su libro dice: «Los hombres tienen derechos, y hay cosas que nadie puede hacerles sin violar esos derechos.» Esta afirmación resulta tan evidente a ojos de Nozick que no le parece nece-

sario fundamentarla. Pero la verdad es que está lejos de ser evidente, hasta el punto de que difícilmente pueda ser objeto de un acuerdo espontáneo entre los miembros de una sociedad plural.

Derechos. Hay mucha gente bienintencionada que no cree que los seres humanos tengamos derechos. Este es, por ejemplo, el punto de vista de Alasdair MacIntyre, que dice que la mejor razón para negar la existencia de los derechos es la misma que tenemos para negar la existencia de las brujas y de los unicornios: nadie ha podido aportar nunca una prueba que pueda ser universalmente aceptada (MacIntyre, 1981, pág. 96).

También hay gente que cree que los seres humanos tenemos derechos, pero por razones que no pueden ser empleadas para convencer al conjunto de la sociedad. Supongamos que yo creo que los seres humanos estamos protegidos por los mismos derechos porque todos somos hijos de Dios. Este es un argumento que puede ser muy fuerte a mis ojos, pero no puedo pretender que sea igualmente aceptado por todos mis conciudadanos.

Por último, hay personas que aceptan la existencia de los derechos pero no los ven como preexistentes al Estado. Antes del Estado no hay derechos sino posibilidades de acción, y el problema es que en el estado de naturaleza nada nos asegura que podamos ejercerlas.

Si los hombres nos damos instituciones políticas no es entonces para proteger ciertos derechos preexistentes, sino pa-

ra proteger ciertas posibilidades de acción que nos parecen moralmente relevantes.

La posibilidad de defender racionalmente cualquiera de estos puntos de vista plantea una seria dificultad para Nozick. Su primer supuesto dice que los hombres tenemos derechos que preexisten al Estado, pero los miembros de una sociedad plural pueden no estar de acuerdo con él. Esta discrepancia tiene consecuencias potencialmente devastadoras sobre su argumentación, ya que no sólo debilita su primer supuesto sino también el segundo.

Regulación. Este segundo supuesto dice que, si tenemos determinado derecho, tenemos derecho a ejercerlo de manera ilimitada. Pero esta afirmación sólo resulta aceptable si antes hemos admitido que los derechos preexisten al Estado. Es evidente que un Estado nacido para reconocer derechos previamente existentes no podría tener ninguna legitimidad para limitarlos. Pero, ¿qué ocurre si una persona no acepta que tengamos derechos? En ese caso, el segundo supuesto de Nozick pierde todo su sentido. ¿Y qué ocurre si alguien piensa que los únicos derechos que tenemos son aquellos que han sido creados por el Estado? En este caso, el segundo supuesto pierde toda su fuerza. Si el Estado es el creador de los derechos, entonces tiene toda la legitimidad necesaria para regular y eventualmente limitar su ejercicio.

La propiedad. Una manera de entender el significado de esta discusión abstracta es aplicarla al caso del derecho de propiedad, que ocupa un lugar crucial en la argumentación de Nozick. El problema planteado por el ejemplo de Wilt Chamberlain proviene de concebir la propiedad como un derecho anterior al Estado y, por tanto, como un derecho que debemos respetar absolutamente. Pero si dejamos de ver a la propiedad como un derecho natural y pasamos a verla como un resultado del contrato social, buena parte del efecto que produce el «ejemplo Chamberlain» tiende a diluirse.

Para poder ver las cosas de esta manera basta con alejarse de Locke (a quien sigue Nozick en su argumentación) y aproximarse a Kant. En la primera parte de su *Metafísica de las costumbres* Kant argumenta que en el estado de naturaleza no hay propiedad sino posesión. Poseemos aquello que somos capaces de defender de la rapacidad ajena. Pero la posesión es una cuestión puramente empírica. Mientras permanezcamos en el estado de naturaleza, un siglo de posesión no me da ningún argumento frente a alguien más fuerte que yo.

Kant explicaba este punto diciendo que la apropiación es un acto unilateral de un individuo respecto de un objeto, mientras que la propiedad es una institución social. Si soy el legítimo propietario de un trozo de tierra (y no simplemente su poseedor) no preciso disponer de la fuerza para defenderlo. El Estado lo hará por mí porque me recono-

ce derechos que no dependen de mi capacidad efectiva de defenderme.

¿Derecho natural? Nozick afirma que si somos legítimos propietarios de algo entonces tenemos derecho a hacer lo que queramos con ese bien sin que nadie pueda imponernos ningún límite. De aquí deduce que todo impuesto redistributivo es ilegítimo. Pero esta idea sólo podría ser vista como evidente si también fuera evidente que la propiedad es un derecho natural. En cambio, si el derecho de propiedad pasa a ser visto como una institución social, entonces las limitaciones que se impongan a su ejercicio no pueden ser consideradas intrínsecamente ilegítimas, esto es, no pueden ser asimiladas al robo. Como observa Thomas Nagel, «el robo viola expectativas individuales legítimas basadas en la institución de la propiedad. Pero no hay *a priori* ninguna expectativa legítima de que el sistema de propiedad elegido por la sociedad no incluya al menos mínimas restricciones a los intercambios voluntarios o transferencias».

Los supuestos de los que parte Nozick están entonces lejos de ser incontrovertidos. Como señala Will Kymlicka, toda su argumentación se limita a decir: «Aquí hay un conjunto de derechos de propiedad absolutos. Respetémoslos y que los individuos los usen como quieran.» Pero lo que falla es la primera parte de la afirmación.

A pesar de esta insuficiencia, la crítica de Nozick tiene la virtud de recordarnos una verdad importante que Philippe

Van Parijs resume de esta manera: discutir «sobre la manera 'óptima' o 'más equitativa' de distribuir niveles de vida puede ser legítimo cuando se trata de repartir un maná que nos cae del cielo. Pero la mayor parte de los componentes de nuestro nivel de vida no cae del cielo. Se trata de objetos —bienes y servicios— que provienen del trabajo y del ahorro de individuos determinados». Este es un problema al que deben prestar mucha atención los liberales que, como Rawls, tienen fuertes pretensiones igualitaristas.

El debate con los igualitaristas

Los libertaristas acusaron a Rawls de no ser suficientemente liberal. Los igualitaristas radicales lo acusaron, en cambio, de no ser suficientemente igualitario. De hecho, ha llegado a afirmarse que, una vez consideradas ciertas dificultades, podríamos descubrir una lectura de Rawls bastante más «a la derecha» de lo que él jamás imaginó. Sin habérselo propuesto, Rawls terminó por descubrir que estaba entre dos fuegos.

La discusión con los igualitaristas es tan compleja como el debate con los utilitaristas. Es también una discusión que se da en varios frentes y a propósito de un conjunto muy heterogéneo de problemas[15]. Pero lo que importa subrayar aquí es que esta discusión puso en evidencia algunas

15 Para una presentación en castellano de algunos de estos desarrollos, *véase* Gargarella, 1999, págs. 69 y sigs.

dificultades que Rawls no había percibido en el momento de redactar *A Theory of Justice*.

Una de esas dificultades tiene que ver con la formulación del *principio de diferencia*. En su formulación original, este principio decía que sólo podían aceptarse aquellas desigualdades que fueran en beneficio de los miembros menos favorecidos de la sociedad. Pero el problema es que esta idea, al menos formulada de este modo, podía conducir a convalidar formas de desigualdad verdaderamente extremas. Para percibir mejor esta dificultad, volvamos una vez más a nuestra sociedad compuesta de dos individuos e imaginemos la siguiente distribución:

A	3
B	4
	7

Como se ve, el individuo A tiene 3 unidades de utilidad y el individuo B tiene 4 unidades. Ahora consideremos la posibilidad de pasar a una situación alternativa planteada en estos términos:

A	4
B	66
	70

Esta situación es mucho más eficiente que la anterior, hasta el punto de que el producto se ha multiplicado por

diez (pasó de 7 a 70). Además se cumple la condición de que el individuo menos favorecido (es decir, el individuo A) ha mejorado su posición. Pero esta mejora es insignificante al lado de los beneficios que obtuvo B. La riqueza total ha aumentado enormemente, pero el individuo A apenas ha podido favorecerse. La sociedad no sólo se ha vuelto más eficiente, sino también muchísimo más desigual. Y sin embargo, parecería que el paso de una situación a otra cumple con la exigencia establecida en el *principio de diferencia*.

Reformulación. La consideración de este problema llevó a Rawls a enfatizar que, para que las desigualdades puedan ser justificadas, no es suficiente con que generen *algún* beneficio para los menos favorecidos. Si nos limitáramos a establecer esta condición, efectivamente estaríamos justificando crecimientos de la desigualdad casi ilimitados. Lo que hay que exigir, por tanto, es que las desigualdades coloquen a los menos favorecidos en la *mejor* situación en la que puedan estar. Esto implica que tenemos que considerar todos los estados del mundo a nuestro alcance y elegir aquel en que los menos favorecidos estén mejor. Esta situación no será seguramente de estricta igualdad, porque ciertas formas de desigualdad pueden tener consecuencias muy favorables en términos de eficiencia. Pero el mismo tiempo es muy probable que nos veamos obligados a rechazar soluciones extremadamente eficientes para asegurar que los menos favorecidos se encuentren en la mejor situación.

Imaginemos que nuestra sociedad ideal de dos individuos sólo tiene a su alcance estos tres estados del mundo:

	I	II	III
A	3	15	4
B	4	35	96
	7	50	100

La situación más eficiente es claramente la tercera, en la que el producto total alcanza 100 unidades. Sin embargo, el *principio de diferencia* nos obligaría a elegir la segunda, que es la situación que más favorece al que está peor. No hay duda de que al hacer esta opción estaremos pagando un alto precio en términos de eficiencia. Pero al mismo tiempo se le estará planteando a la sociedad el desafío de encontrar caminos hacia la eficiencia que sean más compatibles con la exigencia de justicia. Lo que parecía ser un simple criterio para elegir entre posibles pautas de distribución se convierte entonces en una exigencia igualitaria muy severa.

Capacidades. Un segundo frente de discusión con los igualitaristas (igualmente importante para la suerte de la teoría de Rawls) tiene que ver con la definición de *bienes primarios*. Rawls los había definido como «medios para todo propósito» y había afirmado que para cualquier individuo determinado tener más de esos bienes es preferible a

tener menos. Y también había asumido que una mayor disponibilidad de esos bienes iba a favorecer igualmente a todos los miembros de la sociedad. Pero este supuesto fue duramente atacado por algunos filósofos igualitaristas, especialmente el economista indio Amartya Sen (otro premio Nobel de Economía muy involucrado en los debates sobre la igualdad).

Sen mostró que una misma dotación de bienes primarios no tiene necesariamente el mismo impacto sobre la capacidad que tienen diferentes personas de poner en práctica su plan de vida.

Imaginemos dos individuos: uno es una persona normal y el otro sufre un grave impedimento físico que sólo puede ser superado con grandes gastos médicos (por ejemplo, comprando medicamentos muy caros). Si ambos reciben la misma dotación de bienes primarios, el resultado sobre sus vidas será muy diferente: el primero podrá invertir la totalidad de esos bienes en la ejecución del plan que ha elegido, mientras que el segundo deberá invertir la mayor parte en el esfuerzo que necesita para llegar a la situación que para el otro es el punto de partida. Esto creará una diferencia difícilmente defendible en términos de una teoría de la justicia, ya que la desigualdad no resulta de las decisiones tomadas por cada uno de los individuos, sino de una circunstancia que actúa antes de que alguno de ellos haya tomado una decisión.

La respuesta de Rawls a esta objeción no fue demasiado convincente: básicamente consistió en decir que su teoría

está pensada para los «casos normales», y que las situaciones especiales deben ser objeto de un tratamiento *ad hoc*. Lo poco satisfactorio de esta respuesta llevó a Sen a desarrollar un teoría de la justicia alternativa.

Los méritos de esta nueva teoría son un tema en discusión, pero la propia existencia del debate es reveladora del lugar central que ha pasado a ocupar la doctrina de los *bienes primarios* (junto a ideas como la de *velo de ignorancia* o *principio de diferencia)* en la discusión contemporánea sobre la justicia.

El debate con los comunitaristas

El Rawls de los años 70 creía que sus principios de justicia podían ser aceptados por todos los miembros de una sociedad plural, dado que habían sido elegidos mediante un procedimiento que todos podían considerar como válido. Lo esencial de este procedimiento consistía en someterse a las condiciones de la *posición original*.

¿Por qué estas condiciones podían ser vistas como aceptables por todos? Rawls había dado dos respuestas a esta pregunta. La primera era que esas condiciones serían las elegidas por todo individuo racional y razonable. La segunda era que esas condiciones serían las elegidas por todo individuo capaz de adherirse a lo que él llamaba *una concepción delgada (o restringida) del bien,* es decir, una teoría mínima acerca del bien individual, compatible con toda teoría comprehensiva razonable.

Kantismo. La primera respuesta ponía de manifiesto el carácter kantiano de la teoría. Rawls proponía un retrato mínimo del agente moral que en cierto modo desempeñaba el papel de sujeto trascendental. Este retrato presentaba a cada miembro de la sociedad como dotado de las dos capacidades morales fundamentales: la capacidad de elaborar y de intentar poner en práctica una concepción del bien, y la capacidad de desarrollar un sentido de la justicia (es decir, de elegir principios de justicia y de respetarlos). Rawls pensaba que todos los miembros de la sociedad deberían admitir su interés en preservar esas capacidades. Dicho de otro modo: cada miembro de la sociedad debería admitir su preferencia de vivir en una sociedad donde el ejercicio de estas capacidades *no* se vea sistemáticamente amenazado. Y si uno está dispuesto a aceptar este dato, entonces también debería considerar aceptables las restricciones de la *posición original*.

Primera respuesta a la pregunta sobre la justificación del procedimiento, pero claro no era una respuesta suficiente. Si los individuos sólo compartieran su interés en proteger sus *capacidades morales fundamentales,* la elaboración de acuerdos en la *posición original* no podría ir demasiado lejos. Por esta razón Rawls asume que los miembros de la sociedad no sólo comparten un mismo interés en proteger las *capacidades morales fundamentales*, sino también una teoría mínima acerca de lo que es bueno para ellos[16].

16 Para todo lo que sigue, *véase* Rawls, 1971, págs. 395 y sigs., esp., págs. 437 y sigs.

La teoría «delgada» del bien propuesta por Rawls puede resumirse en tres afirmaciones: (I) es mejor poder realizar la propia concepción del bien que no poder realizarla; (II) es preferible disponer de más medios que de menos medios para realizar nuestra concepción del bien; y (III) es posible identificar un conjunto de bienes que sirven como medios para la realización de cualquier concepción del bien razonable (los *bienes primarios).* El supuesto era que los individuos racionales y razonables pueden ponerse de acuerdo en estos tres puntos, a pesar de las diferencias entre sus concepciones del bien. Por tanto, la necesidad de definir una teoría mínima del bien puede satisfacerse sin violar las condiciones impuestas por el hecho del pluralismo.

Rawls resumía esta idea diciendo que su teoría respetaba «la prioridad de lo justo sobre lo bueno». Con esto creía haber resuelto el problema, hasta el punto de que no parecía esperar que su teoría fuera atacada en este punto. Pero esto fue exactamente lo que pasó. Y los responsables del ataque fueron un conjunto de filósofos a los que se conoce con el nombre de *comunitaristas*[17].

¿Qué sujeto? La primera objeción importante fue planteada por Michael Sandel. No es verdad, dice Sandel, que la teoría rawlsiana de la justicia esté fundada en una concepción incontrovertida del agente moral. Los agentes rawlsianos llegan a la *posición original* libres de concepcio-

---

17 Esta discusión está descrita con mayor detalle en Da Silveira, 1994.

nes del bien y libres de antecedentes históricos, pero no libres de todo supuesto. Para que la teoría se sostenga, algunas cosas tienen que ser ciertas acerca de nosotros mismos.

¿Qué es lo que tiene que ser cierto? Básicamente, que seamos individuos libres de todo compromiso con nuestro propio contexto, es decir, individuos capaces de llevar adelante su vida moral manteniendo una distancia respecto de la circunstancia. Ciertamente vamos a ser condicionados por el entorno cultural y social en el que crecemos y actuamos, pero siempre mantendremos la capacidad constitutiva de despegarnos de ese contexto y de tomar decisiones en función de criterios fijados por nosotros mismos.

Sandel llamó a esta descripción la doctrina del «sujeto libre de compromisos» *(unencumbered self)*. Y junto a otros comunitaristas se dedicó a argumentar que en esa doctrina se esconden dos tesis inaceptables. La primera dice que nuestra capacidad de actuar como agentes morales puede ejercerse sin ningún compromiso con una concepción específica del bien. La segunda afirma que la pregunta «¿qué debo hacer?» puede contestarse con independencia de la pregunta «¿quién soy?»

Los comunitaristas se opusieron a la primera afirmación diciendo que hay compromisos y propósitos que no son objeto de ninguna elección, sino que están ligados desde siempre a nuestra biografía moral. Nacemos miembros de ciertas comunidades (nacionales, religiosas, lingüísticas, culturales) con las que nos descubrimos involucrados. Y es en esas comunidades donde encontramos los modelos y el

impulso necesario para constituirnos en agentes morales bien desarrollados. Nadie puede madurar su capacidad de juicio, comprender la idea de virtud ni consolidar una identidad moral, a menos que crezca en una comunidad construida en torno a una concepción sustancial del bien. La promesa de una independencia absoluta es, por tanto, una falsa promesa de libertad: «Imaginar a alguien incapaz de lazos constitutivos [...] no es concebir el ideal de un agente libre y racional, sino imaginar a una persona enteramente sin carácter, sin profundidad moral.»

La biografía moral. Los comunitaristas se opusieron también a la segunda afirmación diciendo que, además de necesitar lazos constitutivos para poder desarrollar nuestra identidad moral, también necesitamos preguntarnos sobre nuestra identidad para darle un sentido a nuestros actos. Una vez que hemos llegado a ser agentes morales maduros, nuestra tarea no consiste solamente en encontrar la respuesta correcta a cada situación específica, sino en conseguir que esas respuestas formen parte de una historia que sea inteligible para nuestra comunidad y para nosotros mismos. Rawls parece creer que «el agente se vincula con sus fines como un sujeto volitivo a los objetos de su elección», pero los comunitaristas afirman que el vínculo entre el sujeto y sus fines es materia de autodescubrimiento.

Teoría del bien. El primer ataque comunitarista se centraba en la teoría implícita del agente moral. Pero hubo todavía un segundo ataque, que se centró en lo que Sandel llama «el problema de la delgadez de la teoría delgada del bien». Esta nueva objeción dice que, por más delgada que sea la teoría «delgada» del bien, su supuesto básico consiste en identificar al bien individual con el libre ejercicio de la capacidad de elegir. Y el problema es que ese supuesto es inaceptable.

Si el modelo del agente moral libre de compromisos fuera admisible, no habría mayores inconvenientes en identificar a nuestro propio bien con la plena realización de aquello que hemos elegido. Pero, dado que no es posible vivir una vida moral desvinculada de un esfuerzo de autodescubrimiento, el bien no puede ser siempre un objeto de libre elección. Al menos parcialmente, nuestro bien consistirá en la plena realización de aquellos vínculos (familiares, nacionales, religiosos) en los que estamos involucrados desde siempre. No existe algo así como una absoluta libertad de elección: no tenemos la libertad de renegar de nuestro pasado ni de ignorar los afectos y compromisos que han sido inseparables del proceso de constitución de nuestra identidad. Podemos ciertamente romper con ellos, pero esa es también una manera de reconocerlos. La conclusión es que, al ignorar estas restricciones e identificar al bien con la realización de aquello que hemos elegido libremente, la teoría del bien implícita en el pensamiento rawlsiano se hace tan discutible como cualquier otra.

¿Principios justos? Si esta crítica fuera correcta, entonces el programa rawlsiano habría recibido un duro golpe. Rawls no sólo había buscado justificar las condiciones más adecuadas para que cada individuo pudiera desarrollar su búsqueda del bien, sino que había pretendido hacerlo sin tomar partido a favor ni en contra de las múltiples concepciones del bien que pueden ser elegidas por los individuos. Quienes discutieron con él durante los años setenta y principios de los ochenta se habían centrado en el primer propósito, pero los comunitaristas estaban atacando ahora la segunda pretensión. Lo que se preguntaban no era si los principios de Rawls eran o no suficientemente justos, sino si eran justos para *todos* los miembros de la sociedad. Y la respuesta que daban era negativa.

Este ataque tuvo consecuencias importantes. En primer lugar, fue la crítica que condujo a Rawls a introducir mayores modificaciones en su teoría. En segundo lugar, puso en marcha una discusión extremadamente célebre que se conoce con el nombre de «el debate entre liberales y comunitaristas». Y el solo hecho de que Rawls haya estado involucrado en esta discusión hizo que su influencia se extendiera bastante más allá de los límites del mundo anglosajón.

El debate entre liberales y comunitaristas se caracterizó por tener una relación compleja con el espacio y con el tiempo. Sin duda fue iniciado por un pequeño núcleo de fi-

lósofos anglosajones contemporáneos, pero a la vez reeditó un conjunto de confrontaciones y de puntos de vista que habían sido formulados en diferentes lugares y momentos. Si tuviéramos que hacer una lista exhaustiva de los «participantes» en el debate deberíamos incluir a varios pensadores de la generación de posguerra (Hannah Arendt y Jean-Paul Sartre, por ejemplo), a autores del siglo pasado como Jeremy Bentham y John Stuart Mill, a filósofos morales y políticos de los siglos XVII y XVIII como Hobbes, Constant o Burke, y a varios de los grandes clásicos de la tradición filosófica (Platón, Aristóteles, Kant, Hegel). Estas múltiples referencias permiten entender el interés que despertó fuera del contexto anglosajón. La discusión entre Rawls y los utilitaristas era casi incomprensible (y sin duda poco interesante) para un francés o un alemán. En cambio, un debate en el que se hablaba del pensamiento político de Aristóteles o de la teoría moral de Kant resultaba muy familiar para los continentales. Al haber generado la reacción comunitarista, y al haberse involucrado en la discusión con quienes la defienden, Rawls contribuyó a debilitar las «fronteras filosóficas» y extendió su influencia mucho más allá de los límites de su mundo de origen.

CAPÍTULO 3

¿Un nuevo Rawls?

Teoría de la justicia se había caracterizado por su fuerte pretensión universalista. Su ambición era establecer un conjunto de principios de justicia cuya validez no dependiera de la diversidad de los contextos en los que fueran a ser aplicados. Rawls había sido muy claro al expresar este propósito. De acuerdo con sus propias palabras, la ficción de la *posición original* nos obligaba a «contemplar la situación humana, no sólo desde todos los puntos de vista sociales sino también desde todos los puntos de vista temporales». En algún momento llegó a hablar del intento de desarrollar una «geometría moral». Estas afirmaciones llevaron a muchos comentaristas a verle como un continuador del pensamiento kantiano más ortodoxo[18].

---

18 Véase, por ejemplo, Höffe, 1988, pág. 63.

En realidad, las cosas fueron más complejas desde el principio. Así como es posible encontrar en *Teoría* muchos elementos típicamente kantianos, también hay puntos de ruptura evidentes[19]. Pero lo que importa aquí es subrayar que, aun si es posible detectar diferencias entre el pensamiento de Rawls y el de Kant, ambos autores tenían en común dos rasgos esenciales: la pretensión de establecer principios normativos cuya validez no fuera puesta en cuestión por la diversidad de los contextos de aplicación, y el intento de fundar esa universalidad en ciertas características constitutivas de los agentes morales.

Este kantismo original fue el blanco de las críticas comunitaristas. En opinión de esos autores, el pensamiento político de Rawls estaba asociado a una tradición tan necesitada de justificación como cualquier otra, y llevaba implícita una psicología moral que no era la de todo agente capaz de responsabilizarse de sus acciones, sino la de aquellos agentes que colocan a la libre elección como supremo valor moral. Bajo un manto de universalismo y de trascendentalismo, había la defensa de un particularismo entre otros.

Rawls fue muy sensible a estas críticas, y el modo en que intentó responderlas lo hizo alejarse de su perspectiva inicial. Esta es una transformación radical de sus ideas, pero una transformación difícil de describir y de evaluar. La ruptura de Rawls con Kant fue un largo proceso, durante

---

19 Para una discusión más detallada, véase Da Silveira, 1997.

buena parte del cual el propio Rawls no parece haber sido enteramente consciente del alcance de las variantes que iba introduciendo.

De manera muy esquemática puede decirse que el Rawls posterior a *Teoría* pasó por dos momentos. El primero puede fecharse en 1980 y consistió en una defensa de lo que él mismo llamó un *constructivismo kantiano*. Se trata de una etapa de transición en la que su pensamiento experimenta modificaciones importantes pero carentes de estabilidad. El segundo momento empieza a tomar forma en la segunda mitad de esa misma década y se consolida en 1993, con la publicación de *El liberalismo político.* En esta nueva etapa, Rawls abandona el *constructivismo kantiano* en favor de lo que llama un *constructivismo político.*

La idea de constructivismo kantiano

En el año 1980 Rawls fue invitado a dictar las conferencias John Dewey, organizadas por la Universidad de Columbia. Las *Dewey Lectures* se realizaron en abril de ese año y en setiembre fueron publicadas con el título: «Constructivismo kantiano en la teoría moral.» Según su propia declaración, lo que Rawls se proponía era «exponer de forma más clara las raíces kantianas» de su teoría[20]. Pero tanto el título como la declaración de propósitos resultaron ser engañosos. En los hechos Rawls estaba rompiendo con el kantismo

20 Cuando se citan textos que fueron recogidos en los *Collected Papers*, se da primero la referencia original y luego la correspondiente a esa edición.

justo en el momento en que parecía declararle su más firme adhesión.

Las *Dewey Lectures* se presentan como una aclaración metodológica acerca de cómo se establece la conexión entre una concepción de la persona y los principios de justicia. El sustantivo «constructivismo» significa que «fuera del procedimiento de construir los principios no hay razones de justicia», es decir, que en la construcción de los principios compartidos no se apela a ningún orden objetivo independiente. Este constructivismo es «kantiano» porque la concepción de la persona en la que se funda es la misma adoptada por Kant: los individuos son vistos como racionales y razonables, mutuamente desinteresados e independientes de sus fines.

Rawls sigue asociando su teoría de la justicia a una concepción de la persona que recoge los principales rasgos atribuidos por Kant al agente moral. Pero un examen más detenido muestra que el modo en que se recoge esta concepción de la persona ha cambiado. La teoría de 1971 era kantiana porque Rawls pensaba que Kant había hecho el mejor catálogo disponible de las características constitutivas del agente moral. Su kantismo consistía en aceptar esa descripción. En cambio, la concepción de la persona que Rawls defiende en las *Dewey Lectures* no surge de ningún análisis filosófico, sino que es vista como implícita en la cultura pública de la sociedad. La teoría sigue siendo kantiana, pero sólo porque la concepción de la persona que está implícita en la cultura pública de las sociedades libe-

ral-democráticas coincide en términos generales con la concepción de la persona delineada por Kant. Lo que justifica una concepción de la justicia es «su congruencia con nuestra más profunda comprensión de nosotros mismos y de nuestras aspiraciones, y el percatarnos de que, dada nuestra historia y las tradiciones inscritas en nuestra vida pública, es la doctrina más razonable para nosotros».

La pretensión transcendental ha sido sustituida por la referencia a una concepción de la persona que aparece ligada a una tradición política específica, y esto tiene como resultado el abandono de toda pretensión de universalidad.

Aunque Rawls sigue diciendo que su teoría es kantiana, el adjetivo «kantiano» ha cambiado entonces de significado. Para el Rawls de 1971, su propio kantismo residía en su pretensión universalista y en la apelación a ciertas características constitutivas de la agencia moral. Para el Rawls de 1980, su kantismo consiste en la adhesión a una concepción de la persona que es vista como implícita en la cultura pública de una sociedad que intenta organizarse políticamente. Los contenidos de esta concepción de la persona coinciden con los de Kant, pero el modo en que Rawls llega a ellos no tiene nada de kantiano. La teoría rawlsiana de la justicia ya no le habla a la totalidad de los individuos racionales y razonables, sino a aquellos identificados con una cultura pública característica de las sociedades liberal-democráticas.

Del constructivismo kantiano al constructivismo político

Las *Dewey Lectures* fueron un intento de reformulación que el propio Rawls no tardó en juzgar insuficiente. Desde ese momento y hasta la aparición de *El liberalismo político* su teoría fue experimentando una serie de modificaciones que en conjunto arrojaron un resultado muy diferente al de 1971.

Aunque las cosas hayan ocurrido de una manera más compleja, la evolución de la argumentación rawlsiana puede resumirse del siguiente modo: para responder a la crítica que le acusaba de haberse comprometido con una reconstrucción discutible de nuestra psicología moral, Rawls introdujo una distinción entre *liberalismo comprehensivo* y *liberalismo político* que profundizaba el movimiento iniciado en las *Dewey Lectures*. Y para responder a la crítica que le acusaba de haberse comprometido con una teoría del bien específica (o al menos con una metateoría controvertida), sustituyó la versión original de su teoría «delgada» del bien por una versión más compleja, que incluye las nociones de *consenso por superposición* y de *cultura política pública*.

El esfuerzo por alejarse de todo riesgo de imponer una reconstrucción discutible de nuestra constitución moral le llevó a enfatizar su ruptura con toda pretensión objetivista. Para Rawls es cada vez más decisivo afirmar que la única forma de objetividad moral a la que podemos aspirar consiste en la construcción de un punto de vista que todos

puedan aceptar. Dicho de otro modo: algo es moralmente objetivo cuando puede ser objeto de un acuerdo razonable.

Razonabilidad. El problema, naturalmente, consiste en definir esta «razonabilidad». Kant había dado una respuesta a esta dificultad, pero Rawls admite ahora que esa solución dependía de nociones que pueden ser objeto de desacuerdo entre los ciudadanos razonables de una sociedad plural. Por tanto, hace falta abandonar el *constructivismo kantiano* para desarrollar la perspectiva constructivista en términos estrictamente políticos. Se trata de un constructivismo caracterizado por tres rasgos.

En primer lugar, el *constructivismo político* no se aplica a la vida moral en su conjunto sino exclusivamente a la vida política, es decir, a aquel tipo de vida en el que debemos justificar normas o acciones específicas de acuerdo a criterios que todos los miembros de la sociedad podamos aceptar. En segundo lugar, el *constructivismo político* no aspira a formular principios universalmente válidos, sino los principios más adecuados para un tipo de sociedad específico, a saber, aquellas caracterizadas por lo que Rawls llama ahora «el hecho del pluralismo razonable». En tercer lugar, y como consecuencia de lo anterior, el *constructivismo político* no apela a una concepción abstracta del agente moral sino a una concepción de la persona «en tanto miembro de una sociedad política». Estas ideas de persona y de sociedad son vistas como componentes fundamentales de la cultura política pública compartida por los ciudadanos.

La objetividad en términos de *constructivismo político* queda entonces definida del siguiente modo: «Decir que una convicción política es objetiva es decir que hay razones [...] suficientes para convencer a todas las personas razonables de que esa convicción es razonable.» Esta definición plantea ciertas dificultades, pero lo que aquí importa es que, al hablar de «personas razonables», Rawls no está hablando de agentes morales kantianos sino de ciudadanos que se reconocen mutuamente como tales. Rawls es todo lo explícito que puede exigirse en este punto: «El agente moral es aquí el ciudadano libre e igual en tanto que es miembro de la sociedad, no el agente moral en general.»

Liberalismo. El tipo de objetividad que nos proporciona el *constructivismo político,* dice Rawls, nos permite defender un liberalismo que ya no será *comprehensivo* sino estrictamente *político.* Un *liberalismo comprehensivo* es un liberalismo que (como los defendidos por Kant o Stuart Mill) se apoya en doctrinas morales o antropológicas «profundas». Tales doctrinas dan por verdaderas algunas tesis extremadamente fuertes a propósito de la naturaleza humana y se comprometen con valores muy específicos sobre lo que da valor a la vida. Los comunitaristas tienen razón al observar que, en la medida en que se asocia a tales ideas y valores, «el liberalismo no es más que otra doctrina sectaria».

Pero el liberalismo rawlsiano no pretende ser comprehensivo sino *político.* La justificación que lo sostiene no está construida sobre una concepción de la naturaleza huma-

na, sino «en términos de ciertos ideales intuitivos, vistos como latentes en la cultura política pública de una sociedad democrática».

Liberalismo comprehensivo y liberalismo político. Esta distinción le permite a Rawls escapar a la objeción de haberse apoyado en una reconstrucción controvertida de nuestra psicología moral. En primer lugar, la cuestión de saber si tenemos o no necesidad de vínculos constitutivos puede quedar abierta. Este problema tiene que ver con la «identidad no pública» de los individuos y no con su «identidad pública», es decir, con la identidad que tenemos cuando se nos mira desde las instituciones. En segundo lugar, la distinción permite rechazar todo vínculo entre la teoría de la justicia y una antropología individualista[21]. Ni Rawls ni ninguno de los liberales más o menos rawlsianos creen que el hombre sea un ser autárquico. Casi todos ellos defienden posiciones solidaristas y les importa señalar que el liberalismo es compatible con ellas.

Pero también les importa dejar claro que, a la hora de diseñar las instituciones y de tomar decisiones políticas, tenemos que respetar la libertad de cada individuo de elegir cómo quiere vivir. El individualismo político dice que, a la hora de distribuir las libertades, la unidad que cuenta es el individuo y no la etnia, la nación no la comunidad. En este

21 Una idea que Rawls ya afirmaba en Rawls, 1972, pág.557. *Véase* también 1971, pág. 584 y 1975, págs. 544 y 550.

sentido estricto, el liberalismo está obligado a ser individualista. Pero toda adhesión a una forma más radical de individualismo debería ser defendida por razones complementarias.

El Rawls de *El liberalismo político* está entonces menos expuesto que el de *Teoría* a la acusación de haberse comprometido con una psicología moral controvertida. Su teoría ya no exige otros presupuestos que los necesarios para formular un concepto normativo de ciudadanía. Las complejidades de la vida moral quedan a un lado. Pero importa observar que la construcción de esta posición más confortable le ha exigido cambios nada despreciables en al menos tres niveles.

Cambios. En primer lugar, las dos *capacidades morales fundamentales* ya no son vistas como características que todos los seres humanos compartimos en tanto que agentes morales, sino como los componentes fundamentales de una concepción del ciudadano justificada en términos políticos. Seguimos estando ante un intento de reconstrucción de nuestra psicología moral, pero esa reconstrucción sólo aspira a reflejar «el modo en que los ciudadanos piensan acerca de sí mismos cuando se plantean problemas de justicia política».

En segundo lugar, este cambio de perspectiva afecta la doctrina de los *bienes primarios*. Estos bienes ya no son vistos como «cosas que se presume que todo ser racional desea», ni como «medios para todo propósito» cuyo catálogo se construye a partir de un examen de los intereses fundamentales de la persona moral kantiana, sino como medios

para satisfacer las necesidades que tienen las personas «como ciudadanos».

Equilibrio reflexivo. Probablemente el cambio más importante sea el que se produce en relación a esta práctica. El Rawls de *Teoría* hablaba de «una sociedad de personas morales», es decir, de una asociación de agentes autónomos que se ven mutuamente como fines en sí mismos. Al elegir los principios de justicia atendiendo exclusivamente a esta constitución de base, los individuos no estaban influidos «por la tradición ni por la autoridad, ni por las opiniones de los otros».

Pero el Rawls de *El liberalismo político* ya no ve a la sociedad como una asociación de agentes morales autónomos, sino como una asociación de ciudadanos que comparten una tradición política específica (aquella que se caracteriza por ver a los ciudadanos como agentes morales autónomos). Por tanto, la elaboración de los principios se vuelve mucho menos individualista. Por una parte, las condiciones de la *posición original* ya no reflejan nuestra naturaleza moral, sino una concepción de la agencia moral contenida en una tradición política específica. Por otra parte, nuestros *juicios bien ponderados* ya no son nuestras intuiciones individuales sino intuiciones públicas, es decir, convicciones que forman parte de las interpretaciones compartidas sobre las tradiciones y las instituciones políticas[22].

---

22 Véase en especial Rawls, 1980, y *Political Liberalism*, pág.45.

Cultura política pública y consenso por superposición

Estos cambios son importantes, pero sólo permiten dar respuesta a la primera crítica comunitarista. Todavía queda por responder la segunda, es decir, aquella que acusa a Rawls de haberse comprometido con una teoría implícita (o al menos una metateoría) del bien individual. Este problema no es menor porque Rawls sigue admitiendo que una teoría de la justicia y una concepción mínima del bien se necesitan mutuamente. Entonces, la pregunta que se plantea es cómo establecer el vínculo entre las exigencias de justicia y las ideas sobre el bien una vez que se ha renunciado al recurso kantiano de apelar a las características propias del agente moral. La nueva respuesta de Rawls consiste en decir que debemos buscarlo en un ámbito específico, que es el ámbito de la *cultura política pública*. Es allí y no en otra parte donde se hace posible construir lo que Rawls llama ahora un *consenso por superposición* o, como se traduce a veces, un *consenso translapado*[23].

Consenso. El *consenso por superposición (overlapping consensus)* es un acuerdo caracterizado por tres rasgos. En primer lugar, se trata de un consenso que tiene por único objeto una

23 En 1989, pág. 233, CP 473, Rawls afirma que la idea de *overlapping consensus* fue introducida en las páginas 387-388 de la edición inglesa de *Teoría* (correspondiente a las páginas 429-430 de la versión castellana). Pero esta afirmación es una generosa interpretación *ex-post*. La realidad es que la idea de consenso por superposición no desempeñaba ninguna función relevante en la versión de 1971.

concepción política de la justicia; las partes se ponen de acuerdo en torno a los principios de justicia que van a regular la convivencia, y sobre ninguna cosa más. En segundo lugar, no se trata de un consenso entre individuos sino entre concepciones comprehensivas razonables (*a grosso modo*, una concepción es razonable si el hecho de que uno o más individuos la ponen en práctica no pone en peligro la capacidad de los otros de poner en práctica sus propias concepciones comprehensivas). Por último, se trata de un consenso que no sólo puede ser suscrito por una diversidad de concepciones comprehensivas razonables, sino al que cada una de ellas llega por sus propias razones.

Rawls presenta un modelo ideal para ilustrar el alcance de esta idea. Se trata de una sociedad hipotética en la que encontramos tres concepciones comprehensivas razonables. La primera es una concepción de inspiración religiosa que incluye las ideas de tolerancia y libertad de conciencia. La segunda es una doctrina liberal fundada en supuestos antropológicos importantes como los que encontramos en Kant o en Mill. La tercera es una concepción pluralista que incluye un conjunto de valores políticos y no políticos vinculados con escasa sistematicidad.

¿Cómo es posible llegar a un consenso en torno a una concepción política de la justicia entre estas diferentes concepciones? Rawls afirma que hay tres maneras de hacerlo y sugiere que existe una tendencia histórica que conduce de la primera a la tercera.

Acuerdo. Una primera manera es concebir el acuerdo en términos de un *modus vivendi*. Se trata del tipo de arreglo al que llegan dos Estados cuando encuentran un punto de equilibrio entre sus intereses. Pero si bien este tipo de transacción es útil a nivel de las relaciones internacionales, resulta insuficiente cuando se trata de justificar instituciones dentro de una sociedad. Un *modus vivendi* sólo puede durar mientras dure la correlación de fuerzas que existía en el momento de pactarlo[24].

Una segunda posibilidad consiste en alcanzar lo que Rawls llama un *consenso constitucional*[25]. Se trata de un consenso que no es «profundo» (se limita a establecer ciertas libertades y derechos sin intentar justificarlos, así como ciertos procedimientos para la toma de decisiones colectivas), y se trata también de un consenso «estrecho»: su campo de aplicación no es la estructura básica de la sociedad sino los procedimientos del gobierno democrático.

Este segundo tipo de consenso representa un progreso respecto al *modus vivendi*, pero todavía tiene debilidades importantes. En primer lugar, el *consenso constitucional* deja demasiadas cosas sin resolver. No especifica el alcance que debe darse a los derechos políticos ni explica qué hacer en caso de que haya conflicto entre ellos. Tampoco dice nada acerca de la igualdad de oportunidades ni de recursos. En segundo lugar, al no contar con ninguna justificación que lo vincule con las

---

24 Esta es la crítica clásica contra las estrategias de inspiración hobbesiana. Véase, por ejemplo, Larmore, 1990, págs. 346 y 359, y Nagel, 1991, págs. 34-35.
25 Véase al respecto Baier, 1989.

concepciones comprehensivas presentes en la sociedad, el *consenso constitucional* no puede asegurar su propia estabilidad[26]

Queda finalmente el *consenso por superposición*. A diferencia del *modus vivendi*, este consenso es moral tanto en su objeto como en su justificación. Su objeto es una concepción política de la justicia que pueda ser justificada mediante argumentos que apelen a ciertas ideas normativas sobre la sociedad y la persona humana. Pero el *consenso por superposición* también se distingue del *consenso constitucional*. En primer lugar, es más «profundo». Los ciudadanos no sólo deben ponerse de acuerdo sobre un conjunto de principios y procedimientos, sino sobre una concepción de la sociedad, sobre una concepción del agente moral, y sobre un conjunto de virtudes públicas asociadas a las ideas de cooperación y de equidad. En segundo lugar, el *consenso por superposición* tiene un mayor alcance que el *consenso constitucional:* la concepción política de la justicia que es su objeto no se aplica exclusivamente a los fundamentos constitucionales sino a toda la estructura básica de la sociedad.

Condiciones. Cuando los miembros de una sociedad llegan a suscribir un *consenso por superposición* se han alejado significativamente de un pacto de no agresión. En primer lugar, la concepción política de la justicia a la que se han adherido no se modifica porque varíe la correlación de fuerzas en

---

26 Rawls agrega todavía otra razón para justificar la insuficiencia de este consenso «estrecho» en *Political Liberalism*, pág. 165.

la sociedad. En segundo lugar, esta concepción no se presenta como la única posible, sino como una concepción moral sostenida por una justificación pública que cada uno puede aceptar por sus propias razones. Por último, esta justificación proporciona argumentos que permiten interpretar los principios y exigencias impuestos por la propia concepción, así como extender estos principios y exigencias del terreno de las garantías formales al de la justicia distributiva.

Rawls afirma que esta clase de consenso no puede alcanzarse en cualquier sociedad. De hecho, sólo es posible en una sociedad plural que ha tenido ciertas experiencias históricas, como la de verificar que el acuerdo unánime en torno a una única concepción comprehensiva sólo es posible como resultado del uso opresivo del poder. También debe ser una sociedad en la que se dan ciertas condiciones razonablemente favorables (económicas, administrativas, tecnológicas), que hagan viable el funcionamiento de las instituciones democráticas. Y debe ser una sociedad en la que una amplia mayoría de ciudadanos (o al menos de ciudadanos políticamente activos) de su adhesión libre y voluntaria a las instituciones democráticas.

En el caso de que estas condiciones se cumplan, será posible embarcarse en la construcción de un *consenso por superposición*. Para hacerlo, cada individuo deberá distinguir entre dos sistemas de valores a los que puede prestar su adhesión: por un lado, la concepción política de la justicia con la que decida comprometerse; por otro, una doctrina

plena o parcialmente comprehensiva con la que se identifique y que de algún modo se vincule con la anterior. La afirmación fundamental del «último» Rawls es que una teoría de la justicia sólo podrá pactarse y ser operativa en el caso de que se logre construir un consenso de esta naturaleza.

Promesas e insuficiencias

Los cambios producidos en el pensamiento de Rawls fueron recibidos muy desigualmente. Muchos de aquellos que habían aplaudido la aparición de *Teoría de la justicia* se declararon decepcionados por el cambio de estilo. El libro de 1971 era (y sigue siendo) una verdadera proeza de elegancia y de economía argumentativa. *El liberalismo político,* en cambio, es una colección de artículos que defienden una posición tal vez demasiado cargada de matices y cualificaciones.

Pero este es una forma de evaluación casi estética a la que no debe prestarse demasiada atención. Una teoría puede ser muy elegante al precio de ser excesivamente simplificadora. Otra puede ser menos límpida pero más rica y poderosa, en el sentido de hacerse más debidamente cargo de las complejidades de la materia considerada. Si este fuera el caso de *El liberalismo político* no habría por qué preocuparse demasiado.

El problema es que no hay acuerdo en este punto. Para aquellos que se identifican con la sensibilidad comunita-

rista, Rawls habría ganado mucho al alejarse del kantismo. Sus ideas serían ahora menos abstractas, más capaces de responder a la variedad de tradiciones y de formas de vida. Para quienes se identifican con la percepción liberal, Rawls habría perdido mucho al romper con el kantismo. Los más duros afirman que simplemente se ha vuelto filosóficamente débil y confuso[27].

Estabilidad

Tal vez el problema más importante en torno al que se discute este punto es lo que se llama el *problema de la estabilidad*[28]. Se trata de saber cuáles son las condiciones que deben cumplirse para que pueda durar un orden político que se ajuste a las exigencias rawlsianas.

Rawls intentó resolver este problema introduciendo las ideas de *overlapping consensus* y de *razón pública*. Su expectativa es que una sociedad que tenga a sus espaldas una saludable tradición democrática optará por esta clase de acuerdo (en lugar del *modus vivendi* o el *consenso constitucional)* y se servirá de él para sustentar una concepción política de la justicia.

¿Qué ocurre si la sociedad no cuenta con los recursos de cultura política que él mismo considera indispensables? En lugar de responder a la pregunta, Rawls avanza la idea optimista de que estas condiciones efectivamente se cumplen

---

27 Bryan Barry ejemplifica este punto de vista.

28 Sobre este problema *véase* Barry, 1995; Da Silveira, 1998, y Gargarella, 1999.

en las sociedades democráticas: «Somos los beneficiarios de tres siglos de pensamiento democrático y de práctica constitucional en desarrollo; y podemos suponer no sólo cierta comprensión de, sino cierta adhesión a, las ideas y valores democráticos tal como se encarnan en las instituciones políticas existentes.» Cuando los ciudadanos inician una deliberación política en este contexto, encuentran que «sus juicios convergen lo suficiente como para que la cooperación política sobre la base del respeto mutuo pueda ser mantenida».

Cuatro concepciones comprehensivas. Todo esto no constituye más que un conjunto de afirmaciones de hecho. ¿Qué ocurre si esos presupuestos empíricos no se cumplen, o sólo se cumplen parcialmente? Para analizar este problema imaginemos una sociedad donde existen cuatro concepciones comprehensivas: la primera incluye una concepción «antropológica» del liberalismo fundada en las ideas de Kant o de Mill. La segunda es una concepción pluralista (hasta aquí repito el modelo de Rawls). La tercera es una concepción hedonista que sólo contiene ideas sobre el bienestar individual. La cuarta es una concepción con la que se identifican los miembros de una minoría étnica e incluye la preferencia en favor de una forma de gobierno jerárquico-patriarcal.

Agreguemos que los adherentes a las dos primeras concepciones no constituyen una clara mayoría en la sociedad, de modo que una concepción política de la justicia sólo po-

drá estabilizarse si es posible construir una justificación pública que recabe la lealtad de al menos parte de quienes prefieren la tercera y la cuarta concepción. ¿Es posible que esto llegue a ocurrir?

Parece claro que éste no es un objetivo que pueda lograrse de manera inmediata. Ni en la concepción hedonista ni en la concepción tribalista se encuentran ideas que lo permitan. La concepción hedonista sólo contiene ideas relativas al bienestar individual. El que la vida en sociedad facilite o dificulte la realización de estas ideas es visto como un dato externo al que debemos adaptar nuestras estrategias personales. En cuanto a la cuarta concepción, las ideas políticas que contiene excluyen la noción de igualdad, de modo que difícilmente se pueda lograr un acuerdo con quienes se identifican con las dos primeras concepciones.

La única manera en que podría vencerse esta dificultad consistiría en embarcarnos en una discusión con los partidarios de la tercera y la cuarta concepción, con el fin de lograr que introduzcan en ellas ciertas modificaciones. Pero esos cambios no pueden ser cosméticos. Para asegurar la estabilidad de una concepción política de la justicia no basta con apelar al interés de nuestros interlocutores, sino que tenemos que conseguir que sean leales al espíritu de las instituciones. Si no alcanzamos este punto, no podremos lograr un *consenso por superposición*.

¿Cuál es el espacio para llevar adelante este intento? Rawls responde que ese espacio es el de la *razón pública*. Pero el modo en que utiliza este concepto es muy diferente de la

concepción tradicional hecha famosa por Kant[29], de manera que hay que empezar por distinguir claramente entre ambas.

Razón. Para Kant hacemos un uso público de la razón cuando nos dirigimos a un público integrado por individuos libres, iguales y dotados de razón, que pueden ser persuadidos mediante la lógica del mejor argumento. El espacio público, entendido como aquel en el que hacemos un uso público de la razón, está caracterizado por dos rasgos: la igualdad (los argumentos valen por su propio peso y no en función de quien los propone) y la inclusividad (puede participar en la discusión todo aquel que tenga algo que decir). El uso público de la razón en el sentido kantiano no depende así del tema en discusión ni de la cantidad de interlocutores directos. Lo único que cuenta es el tipo de argumento que utilizamos.

Razón pública. Rawls, en cambio, establece su definición de *razón pública* en función de quién discute y de cuál es el tema en discusión. Desde su punto de vista, la *razón pública* es «el razonamiento de los ciudadanos en el foro público a propósito de los fundamentos constitucionales y de las cuestiones básicas de justicia».

El primer componente de esta definición es que la *razón pública* es ejercida por los ciudadanos de una democracia

29 Rawls reconoce este punto en *Political Liberalim*, pág. 213n

constitucional. Quienes ejercen esta razón son los ciudadanos cuando participan en campañas políticas y votaciones, los legisladores y agentes del gobierno, los jueces y, muy especialmente, los miembros de la Suprema Corte de Justicia. El segundo elemento de la definición es que solamente algunos de los temas que componen la agenda política caen bajo el dominio de la *razón pública*. Los temas que quedan fuera corresponden a lo que se suele llamar la «política ordinaria» *(ibid)*.

Lo que se opone a la *razón pública,* dice Rawls, no es la razón privada (que en rigor no existe) sino la razón no pública. A este dominio pertenecen nuestras deliberaciones y reflexiones personales, pero también los intercambios de argumentos que se producen en el interior de organizaciones tales como universidades, sociedades científicas, iglesias y grupos profesionales.

Rawls coloca así en el ámbito de la razón no pública algunas formas de argumentación que para Kant eran públicas. Acerca de este tipo de argumentación (por ejemplo, la que se desarrolla en universidades y sociedades científicas) Rawls dice que pertenece a la *cultura de fondo (background culture)* o, como la llama más tarde, la «cultura de la sociedad civil».

Cultura. La *cultura de fondo* es el ámbito donde presentamos y confrontamos nuestras convicciones religiosas, filosóficas y morales. Es, por tanto, el lugar donde reforzamos o modificamos nuestras adhesiones, y donde podemos aspirar a mo-

dificar las concepciones comprehensivas de los demás. Pero para Rawls es vital que esta argumentación no pública se mantenga diferenciada de la discusión a propósito de los fundamentos constitucionales y de las cuestiones básicas de justicia.

El problema queda así planteado en los siguientes términos: el *consenso por superposición* es un acuerdo que se construye en el espacio público. Los argumentos que podemos emplear para construirlo deben satisfacer lo que para Rawls es el criterio identificatorio de las razones públicas; a saber: que puedan ser vistos como parte de una argumentación desarrollada en la Suprema Corte de Justicia. Pero la *cultura de fondo* que hace posible ese consenso pertenece a la órbita no pública y sólo puede ser modificada mediante argumentos no públicos. Cuando entramos en este terreno podemos invocar nuestras convicciones profundas (por ejemplo, religiosas) para explicar por qué la adhesión a esas convicciones nos lleva a ser fieles a los principios de la democracia liberal. Pero el punto es que, si pretendemos que las instituciones políticas funcionen adecuadamente, esta apelación pública a las convicciones privadas no puede ir más allá de una explicitación de nuestras propias motivaciones. Si queremos generar acuerdos políticos estables, tenemos que desarrollar justificaciones públicas en sentido estricto, es decir, justificaciones que *todos* podamos aceptar en tanto que ciudadanos.

La doctrina del *overlapping consensus* muestra así su insuficiencia como respuesta al problema de la estabilidad. Dicho problema consistía en saber cómo una concepción política de la justicia puede generar apoyos duraderos en una sociedad caracterizada por el hecho del pluralismo. La respuesta rawlsiana consiste en decir que puede hacerlo por la vía de generar un *consenso por superposición,* pero ahora sabemos que con esta afirmación no hemos hecho más que postergar el problema: si el *consenso por superposición* es la condición de estabilidad de la concepción de la justicia, queda por saber cómo aseguraremos la existencia de la cultura de fondo que hace posible el *consenso por superposición*. Y Rawls no tiene respuesta para este problema porque lo considera un asunto extra-político.

No hay acuerdo entre los filósofos políticos acerca de si esta manera de plantear las cosas es o no es satisfactoria. Pero el punto es que, a más de treinta años de la publicación de *Teoría de la justicia* el grueso de la discusión sigue girando en torno a las ideas de Rawls. Y este es un mérito del que pocos autores pueden vanagloriarse.

Bibliografía de John Rawls

1951

• «Outline of a Decision Procedure for Ethics». *Philosophical Review* 60/2, págs.177-97. *Collected Papers,* págs. 1-19. Versión castellana: «Esbozo de un procedimiento de decisión para la ética». En John Rawls: *Justicia como equidad. Materiales para una teoría de la Justicia.* (Miguel A. Rodilla, ed.), Madrid, Tecnos, 1986, págs. 1-17.

1955

• «Two Concepts of Rules». *Philosophical Review* 64/1, páginas 3-32. *Collected Papers* 20-46. Versión castellana: «Dos conceptos de reglas». En Philippa Foot (ed.): *Teorías sobre la ética.* México, FCE, 1974, págs. 210-47.

1958

• «Justice as Fairness». *Philosophical Review* 67/2, páginas 164-94. *Collected Papers* 47-72. Versión castellana: «Justicia como equidad». En Rodilla (ed.) 1986, págs. 19-39.

1963

• «Constitutional Liberty and the Concept of Justice». *Nomos* VI, 98-125. *Collected Papers,* págs. 73-95.

• «The Sense of Justice». *Philosophical Review* 72/3, págs.281-305. *Collected Papers* 96-116. Versión castellana: «El sentido de la justicia». En Rodilla (ed.) 1986, págs. 40-57.

1964

• «Legal Obligation and the Duty of Fair Play». En Sidney Hook (ed.): *Law and Philosophy: A Symposium.* Nueva York, New York University

Press, págs. 3-18. *Collected Papers*, págs. 117-29.

1967

• «Distributive Justice». En Peter Laslett & W.G: Runcinam (eds.): *Philosophy, Politics, and Society*. Oxford, Basil Blackwell, págs. 58-82. *Collected Papers* 130-53. Versión castellana: «Justicia distributiva». En Rodilla (ed.), 1986, págs. 58-89.

1968

• «Distributive Justice: Some Addenda». *Natural Law Forum* 13, págs. 51-71. *Collected Papers*, págs. 154-75.

1969

• «The Justification of Civil Disobedience». En Hugo Bedau (ed.): *Civil Disobedience: Theory and Practice*. Nueva York, Pegasus, págs. 240-55. *Collected Papers*, págs. 176-89. Versión castellana: «La justificación de la desobediencia civil». En Rodilla (ed.) 1986, págs. 90-101.

1971

• «Justice as Reciprocity». En Samuel Gorowitz (ed.), *John Stuart Mill: Utilitarianism, with Critical Essays*. Indianápolis, Bobbs-Merrill. *Collected Papers*, págs. 190-224.

• *A Theory of Justice*. Cambridge, Mass., Harvard University Press. Versión española: *Teoría de la justicia*, México, FCE, 1978.

1972

• «Reply to Lyons and Teitelman». *Journal of Philosophy* 69, págs.556-557.

1974

• «Some Reasons for the Maximin Criterion». *American Economic Review* 64, págs. 141-46. *Collected Papers*, págs. 225-31.

• «Reply to Alexander and Musgrave». *Quarterly Journal of Economics* 88, 633-55. *Collected Papers* 232-53. Versión castellana: «Réplica a Alexander y Musgrave». En Rodilla (ed.), 1986, págs. 102-21.

1975

• «A Kantian Conception of Equality». *Cambridge Review* 96, 94-99. Reimpreso bajo el título «A Well-Ordered Society» en Peter Laslett & W.G: Runcinam (eds.): *Philosophy, Politics, and Society*, Fifth Series, Nueva Haven, Yale University Press, 1979. *Collected Papers* 254-66.

• «Fairness to Goodness». *Philosophical Review* 84, págs. 536-554. *Collected Papers* 267-85.

• «The Independence of Moral Theory». *Proceedings and Addresses of the American Philosophical Association* 48, págs. 5-22. *Collected Papers* 286-302. Versión castellana: «La independencia de la teoría moral». En Rodilla (ed.) 1986, págs. 122-136.

1977

• «The Basic Structure as Subject». En Alan Goldman & Jaegwon Kim (eds.): *Values and Morals*. Dordrecht, Reidel, págs. 47-71. Una versión preliminar y más breve había aparecido en *American Philosopical Quarterly* de 14 abril de 1977. Una última versión aparecerá en *Political Liberalism*, págs. 257-88.

1980

• «Kantian Constructivism in Moral Theory». *Journal of Philosophy* 77, 515-72. *Collected Papers* 303-58. Versión castellana: «El constructivismo kantiano en la teoría moral». En Rodilla (ed.) 1986, págs. 137-86.

1982

• «Social Unity and Primary

Goods». En Amartya Sen & Bernard Williams (eds.): *Utilitarianism and Beyond*. Cambridge, Cambridge University Press, págs. 159-186. *Collected Papers* 359-87. Versión castellana: «Unidad social y bienes primarios». En Rodilla (ed.), 1986, págs. 187-211.

1983

• «The Basic Liberties and their Priority». En Sterling McMurrin (ed.): *The Tanner Lectures on Human Values*. Salt Lake City, University of Utah Press, pp. 3-87. Versión castellana: «Sobre las libertades». En John Rawls: *Sobre las libertades*, Barcelona, Paidós, 1990. Traducción de Jorge Vigil Rubio. Este texto será retomado en *Political Liberalism*, págs. 289-371.

1985

• «Justice as Fairness: Political not Metaphysical». *Philosophy & Public Affairs*, 14/3, págs. 223-51. Versión castellana: «Justicia como equidad. Política, no metafísica». En *La Política* (Barcelona) págs., 1, 23-46, traducción de Sebastián Mazzuca. *Collected Papers*, págs. 388-414.

1987

• «Preface» a la edición francesa de *A Theory of Justice*. París, Éditions du Seuil, traducción de Catherine Audard. *Collected Papers*, págs. 415-20.

• «The Idea of an Overlapping Consensus». *Oxford Journal for Legal Studies*, págs. 7/1, 1-25. *Collected Papers*, págs. 421-448.

1988

• «The Priority of Right and Ideas of the Good». *Philosophy & Public Affairs* 17/4, 251-76. *Collected Papers*, págs. 449-72.

1989

• «The Domain of the Politi-

cal and Overlapping Consensus». *New York University Law Review* 64, págs. 2, 233-55. *Collected Papers,* págs. 473-96.

• «Themes in Kant's Moral Philosphy». En Eckart Forster (ed.): *Kant's Trascendental Deductions*. Palo Alto, Stanford University Press, págs. 81-113. *Collected Papers,* págs. 497-528.

1990

• *Justice as Fairness: A Restatement*. Notas de curso de Harvard, inédito.

1993

• *Political Liberalism*. Nueva York, Columbia University Press. Versión castellana: *Liberalismo político*, Barcelona, Crítica, 2001.

• «The Law of Peoples». En S. Shute y S. Hurley (eds.): *On Human Rights: The Oxford Amnesty Lectures, 1993*. Nueva York, Basic Books. *Collected Papers,* págs. 529-64. Traducción al castellano: *El derecho de gentes*. Barcelona, Paidós, 2001.

1995

• «Reply to Habermas». *The Journal of Philosophy* (92/3, págs. 132-80). También incluido en la edición de bolsillo de *Political Liberalism* (1996).

• «Fifty Years After Hiroshima». *Dissent*, verano de 1995, págs. 323-27. *Collected Papers,* págs. 565-73.

1996

• «Introduction to the Paperback Edition» de *Political Liberalism*. Nueva York, Columbia University Press, XXXVI-LXII.

1997

• «The Idea of Public Reason Revisited». *The University of Chicago Law Review,* págs. 64, 3,

765-807. *Collected Papers*, págs. 573-615.

1999

• *Law of Peoples*. Cambridge, Mass. Harvard University Press. (Incluye los artículos «The Law of Peoples» y «The Idea of Public Reason Revisited»). Edición castellana: *El derecho de gentes* y *Una revisión de la idea de razón práctica*. Barcelona, Paidós, 2001.

• *Collected Papers*. Edición de S. Freeman. Cambridge, Mass., Harvard University Press.

2000

• *Lectures on the History of Moral Philosophy*. Edición de Barbara Herman. Cambridge, Mass., Harvard University Press. Edición castellana: *Lecciones sobre la historia de la filosofía moral*. Barcelona, Paidós, 2001.

2001

• *Justice as Fairness. A Restatement*. Edited by Erin Kelly. Cambridge, Mass., Harvard University Press, 2001. Edición castellana: *La justicia como equidad. Una reformulación*. Barcelona, Paidós, 2002.

Otras obras citadas

Audard, C. (ed.)

- 1996: «Utilitarisme». En M. Canto (ed.): *Dictionnaire d'éthique et de philosophie morale.* París, PUF, págs. 1563-1570.

Audard, C.

- 1999: *Anthologie historique et critique de l'utilitarisme.* París, PUF.

Baier, K.

- 1989: «Justice and the Aims of Political Philosophy». *Ethics* 99, págs. 771-790.

Barry, B.

- 1991: *Liberty and Justice. Essays in Political Theory 2*. Oxford, Clarendon Press.
- 1995: «John Rawls and the Search for Stability». *Ethics* 105/4, págs. 874-915.

Berten, A.; Da Silveira, P., y Pourtois, H.

- 1996: *Libéraux et Communautariens*. París, PUF.

Brandt, R.B.

- 1979: *The Good and the Right.* Oxford, Clarendon Press.

Da Silveira, P.

- 1994: «Teoría de la justicia y concepciones del bien: el caso de John Rawls *versus* Sandel & Co.». *Revista Latinoamericana de Filosofía,* (Buenos Aires) 20, 1, págs. 63-86.
- 1997: «Del Rawls kantiano al Rawls humeano: una hipótesis de lectura». *Daimon* (Murcia) 15, págs. 71-81.
- 1998: «La teoría rawlsiana de la estabilidad: consenso por superposición, razón pública y discontinuidad». En F. Quesada (ed.): *La filosofía política en perspectiva.* Barcelona, Anthropos.

Da Silveira, P., y Norman, W.

- 1995:«Rawlsianismo metodológico: una introducción a

la metodología dominante en la filosofía anglosajona contemporánea». *Revista Internacional de Filosofía Política* 5 (Madrid & México), págs. 125-152.

De Marneffe, P.

- 1994: «Rawls's Idea of Public Reason». *Pacific Philosophical Quarterly* 75, págs. 232-50.

Elster, J.

- 1983: *Sour Grapes. Studies in the Subversion of Rationality*. Cambridge, Cambridge University Press.

Fehige, C. y Wessels, U.

- 1998: «Preferences. An Introduction». En C. Fehige y U. Wessels (eds.). *Preferences. Perspectives in Analytical Philosophy*. Berlín/Nueva York, Walter de Gruyter, págs. XX-XLIII.

Gargarella, R.

- 1999: *La justicia después de Rawls*. Barcelona, Paidós.

Goodin, R.

- 1991: «Utility and the Good». En Peter Singer (ed.): *A Companion to Ethics*. Oxford, Basil Blackwell, págs. 241-248.

Harrod, R.F.

- 1936: «Utilitarianism Revisited», *Mind* 45, págs. 137-156.

Harsanyi, J.

- 1975: «Can the Maximin Principle Serve as a Base for Morality? A Critique of John Rawls's Theory». *American Political Science Review* 69, págs. 594-606.
- 1977: *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge, Cambridge University Press.

Höffe, O.

- 1988: «Dans quelle mesure la théorie de John Rawls est-elle kantienne?». En C. Audard *et al.*: *Individu et justice sociale. Au-*

tour de John Rawls. París, Éditions du Seuil, págs. 54-72).

Ingram, A.

• 1994: *A Political Theory of Rights*. Oxford, Oxford University Press.

Kant, I.

• 1783: «Beantwortung der Frage: Was ist Aufklärung?». En Immanuel Kant, *Werke* (Herausgegeben von Wilhelm Weischedel). Darmstadt, Wissenshaftliche Buchgesellshaft, 1968. Band 9, págs. 53-64.

Kymlicka, W.

• 1990: *Contemporary Political Philosophy. An Introduction*. Oxford, Oxford University Press.

Larmore, C.

• 1990: «Political Liberalism». *Political Theory* 18/3, págs. 339-360.

MacIntyre, A.

• 1981: *After Virtue*. Londres, G. Duckworth (2.ª ed., 1985). Versión española: *Tras la virtud*. Barcelona, Crítica, 1987.

• 1984: «Is Patriotism a Virtue?». *The Lindley Lecture*. Kansas, Kansas Univ. Press.

Nagel, T.

• 1975: «Rawls on Justice». En Norman Daniels (ed.): *Reading Rawls Critical Studies on Rawls's Theory of Justice*. Oxford, Basil Blackwell, págs. 1-16.

• 1991: *Equality and Partiality*. Oxford, Oxford University Press.

Nozick, R.

• 1974: *Anarchy, State, and Utopia*. Nueva York, Basic Books.

O'Neill, O.

• 1997: «Political Liberalism and Public Reason: A Critical Notice of John Rawls, Political Liberalism». *The Philosophical Review* 106/3, págs. 411-428.

Sandel, M.

- 1982: *Liberalism and the Limits of Justice.* Cambridge, Cambridge University Press.
- 1984: «The Procedural Republic and the Unencumbered Self». *Political Theory* 12/1, págs. 81-96.
- 1996: *Democracy's Discontent.* Cambridge, Mass. Harvard University Press.

Sen, A.

- 1981: «Rights and Agency». *Philosophy and Public Affairs* II, 1, págs. 3-39.
- 1992: *Inequality Reexamined.* Oxford, Clarendon Press.

Sen, A., y Williams, B. (eds).

- 1982: *Utilitarianism and Beyond.* Cambridge, Cambridge University Press.

Taylor, C.

- 1985: «Atomism». En *Philosophical Papers* Vol. II. Cambridge, Cambridge University Press, págs. 187-210.
- 1988: «Le Juste et le bien». *Revue de Métaphysique et de Morale* 83, págs. 33-55.

Van Parijs, P.

- 1988:«De la eficiencia a la libertad». *Cuadernos del CLAEH* (Montevideo) 47, 105-125.
- 1998: «Pensar la justicia con y contra Rawls». *Cuadernos del CLAEH* (Montevideo) págs. 81-82, 25-33.

Glosario

Bienes primarios *(primary goods)*:

Conjunto de bienes fundamentales que deben ser distribuidos de acuerdo a las exigencias de los principios de justicia. Rawls los definió sucesivamente como «cosas que se presume que todo ser racional desea», como «medios para todo propósito» y como medios para satisfacer las necesidades que tienen las personas «como ciudadanos». El listado de los bienes primarios incluye los derechos y libertades fundamentales, los poderes y prerrogativas asociados a las diferentes funciones y posiciones que podemos ocupar en la sociedad, la riqueza y el ingreso, y las «las bases sociales del respeto de sí mismo».

Capacidades morales fundamentales *(moral powers)*:

Las dos capacidades características de todo agente moral. Se trata de la capacidad de desarrollar y de intentar poner en práctica una concepción del bien (esto es, un conjunto de ideas acerca de cómo vale la pena vivir que orientará sus elecciones particulares) y la capacidad de desarrollar un sentido de la justicia (esto es, la capacidad de entender los principios normativos que se haya decidido respetar y de actuar en consonancia con ellos).

Rawls las llama respectivamente *racionalidad* y *razonabilidad*.

Consenso por superposición *(overlapping consensus)*: Consenso que hace posible la estabilidad de una concepción política de la justicia. Se trata de un consenso que puede ser suscrito por una diversidad de concepciones comprehensivas razonables; cada una de ellas llega a él por sus propias razones.

El objeto del acuerdo son los principios de justicia que regularán la convivencia.

Sin embargo, dado que se trata de un consenso moral, esto requiere también un acuerdo sobre una concepción de la sociedad, sobre una concepción del agente moral, y sobre un conjunto de virtudes públicas asociadas a las ideas de cooperación y de equidad.

Equilibrio reflexivo *(reflective equilibrium)*: Método de validación al que debe ajustarse la reflexión sobre principios de justicia. El equilibrio reflexivo requiere un movimiento de vaivén entre los principios abstractos a los que llegamos mediante el esfuerzo de construcción teórica, y los juicios bien ponderados (o intuiciones morales) que encontramos en nuestra vida moral. El *equilibrio reflexivo* se alcanza cuando los principios y las intuiciones están (al menos provisionalmente) «en línea», es decir, cuando no hay disonancias importantes entre los principios que elegimos y los juicios que han sobrevivido a este proceso de confrontación.

Hecho del pluralismo *(Fact of pluralism)*: Dato fundamental a partir

del cual se debe hacer filosofía política. Se trata de la radical diversidad de convicciones morales, metafísicas y religiosas con las que se identifican los integrantes de las sociedades democráticas contemporáneas. Esta diversidad es el resultado del libre intercambio de ideas y del respeto hacia múltiples formas de vida. No podemos esperar a superarla, por tanto, para embarcarnos en la tarea de construir instituciones comunes. El desafío consiste en diseñar instituciones políticas que puedan ser reconocidas como legítimas en un contexto de diversidad.

Juicios bien ponderados ***(well considered judgements, considered convictions):*** Juicios normativos particulares emitidos bajo condiciones especialmente favorables. Funcionan como «puntos fijos provisionales» de nuestra sensibilidad moral. Esto no significa que sean infalibles, pero sí significa que nos ofrecen el mayor grado de confiabilidad al que podemos aspirar los seres humanos. Estamos seguros de que no hay mayores ausencias de información, que hemos examinado todas las caras del problema y que lo hemos hecho con el deseo sincero de llegar a una conclusión correcta.

Libertades fundamentales ***(basic liberties):*** Libertades que merecen el mayor grado de protección en el orden institucional. No pueden ser sacrificadas por razones de eficiencia ni para asegurar una mejor dis-

tribución de bienes ni oportunidades. Una libertad es fundamental o básica si su restricción o eliminación pone seriamente en peligro la capacidad de ejercer nuestras capacidades morales fundamentales.

Posición original *(original position):*
Asamblea ficticia de la que participan los miembros de la sociedad (directamente o a través de representantes) para definir los principios de justicia y las instituciones sociales fundamentales que condicionarán todo el desarrollo posterior de sus vidas.

Razón pública *(public reason):*
Es la clase de razonamiento en la que se embarcan los ciudadanos de una democracia constitucional cuando discuten en el foro público a propósito de los fundamentos constitucionales y de las cuestiones básicas de justicia. Es típicamente ejercida por los ciudadanos cuando participan en campañas políticas y votaciones, los legisladores y agentes del gobierno, los jueces y, muy especialmente, los miembros de la Suprema Corte de Justicia.

Velo de ignorancia *(veil of ignorance):*
Principal característica de la «posición original». Al ingresar en la asamblea, los individuos «olvidan» cuáles son sus características particulares: no saben si son ricos o pobres, hombres o mujeres, blancos o negros, creyentes o no creyentes. Tampoco conocen cuál es su propia concepción del bien,

es decir, el «sistema de fines» que va a orientar sus elecciones. Esto no implica que los individuos lo olviden todo. Los participantes tienen abundante información general (por ejemplo, saben que hay negros y blancos, mujeres y hombres, creyentes y no creyentes). Lo que desconocen es *su propio lugar* en la sociedad y cuáles son sus ideas acerca de cómo vale la pena vivir.

Rawls en Internet

La más simple búsqueda de materiales de Rawls en Internet amenaza con sepultar a quien consulta bajo decenas de miles de documentos. La gran mayoría de esos materiales está en inglés y se compone de artículos más o menos circunstanciales sobre Rawls y su obra (comentarios a sus diferentes libros, obituarios, etc.).

- http://www.policylibrary.com/ rawls/index.htm. El mejor punto de partida es la **John Rawls Resource Page**, basada en Policy Library. El sitio contiene una bibliografía de Rawls, artículos sobre sus ideas, revisiones bibliográficas, un glosario de expresiones típicas, conexiones para comprar libros y otros *links* de utilidad.
- sun3.lib.uci.edu/~scctr/philosophy/rawls.html. Una bibliografía muy completa hasta el año 1995 puede encontrarse en esta página del departamento de Filosofía de la Universidad de California en Irvine.
- www.ac-toulouse.fr/philosophie/phpes/rawls.htm. Un buen sitio de materiales en francés es la página dedicada a Rawls en el sitio web de la Academia de Toulouse (Francia). Incluye bibliografía de y sobre Rawls, textos escritos por el propio Rawls y por diferentes comentaristas y *links* de utilidad.

- www.eumed.net/cursecon/economistas/Rawls.html. Existen pocos sitios en castellano dedicados a la figura y la obra de Rawls. Una opción es la página dedicada a 250 grandes economistas, a cargo de Juan Martínez Coll, de la Universidad de Málaga. Además de referencias generales se incluye una bibliografía elaborada por Pablo Miró Rocasolano.
- amazon.com. Todos los libros de Rawls están a la venta actualmente, de modo que una visita a una buena librería virtual permite tener una idea precisa de lo que puede adquirirse en forma de libro.

Pablo da Silveira (Montevideo, 1962). Doctor en Filosofía por la Universidad de Lovaina (Bélgica), profesor de Filosofía Política en la Universidad Católica del Uruguay. Ha publicado numerosos artículos en revistas académicas de Europa, América del Sur y del Norte, así como los siguientes libros: *La segunda reforma* (Montevideo, CLAEH/Fundación Banco de Boston, 1995), *Historias de Filósofos* (Buenos Aires, Alfaguara, 1997), *Le débat libéraux-communautariens* (París, Presses Universitaires de France, 1997, en colaboración con André Berten y Hervé Pourtois), *Política & tiempo* (Buenos Aires, Taurus, 2000) y *Diálogo sobre el liberalismo* (Montevideo, Taurus, 2001, en colaboración con Ramón Díaz).